Cochrane Library

Cochrane Database of Systematic Reviews

Psychological therapies for panic disorder with or without agoraphobia in adults: a network meta-analysis (Review)

Pompoli A, Furukawa TA, Imai H, Tajika A, Efthimiou O, Salanti G

Pompoli A, Furukawa TA, Imai H, Tajika A, Efthimiou O, Salanti G.
Psychological therapies for panic disorder with or without agoraphobia in adults: a network meta-analysis.
Cochrane Database of Systematic Reviews 2016, Issue 4. Art. No.: CD011004.
DOI: 10.1002/14651858.CD011004.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Intervention Review]

Psychological therapies for panic disorder with or without agoraphobia in adults: a network meta-analysis

Alessandro Pompoli[1], Toshi A Furukawa[2], Hissei Imai[2], Aran Tajika[2], Orestis Efthimiou[3], Georgia Salanti[4]

[1]Private practice, no academic affiliations, Malcesine, Italy. [2]Department of Health Promotion and Human Behavior, Kyoto University Graduate School of Medicine/School of Public Health, Kyoto, Japan. [3]Department of Hygiene and Epidemiology, University of Ioannina School of Medicine, Ioannina, Greece. [4]Institute of Social and Preventive Medicine (ISPM) & Bern Institute of Primary Care (BIHAM), University of Bern, Bern, Switzerland

Contact: Alessandro Pompoli, Private practice, no academic affiliations, Le grotte 12, Malcesine, Verona, 37018, Italy. alepompoli@msn.com.

Editorial group: Cochrane Common Mental Disorders Group.
Publication status and date: New, published in Issue 4, 2016.

Citation: Pompoli A, Furukawa TA, Imai H, Tajika A, Efthimiou O, Salanti G. Psychological therapies for panic disorder with or without agoraphobia in adults: a network meta-analysis. *Cochrane Database of Systematic Reviews* 2016, Issue 4. Art. No.: CD011004. DOI: 10.1002/14651858.CD011004.pub2.

ABSTRACT

Background

Panic disorder is characterised by the presence of recurrent unexpected panic attacks, discrete periods of fear or anxiety that have a rapid onset and include symptoms such as racing heart, chest pain, sweating and shaking. Panic disorder is common in the general population, with a lifetime prevalence of 1% to 4%. A previous Cochrane meta-analysis suggested that psychological therapy (either alone or combined with pharmacotherapy) can be chosen as a first-line treatment for panic disorder with or without agoraphobia. However, it is not yet clear whether certain psychological therapies can be considered superior to others. In order to answer this question, in this review we performed a network meta-analysis (NMA), in which we compared eight different forms of psychological therapy and three forms of a control condition.

Objectives

To assess the comparative efficacy and acceptability of different psychological therapies and different control conditions for panic disorder, with or without agoraphobia, in adults.

Search methods

We conducted the main searches in the CCDANCTR electronic databases (studies and references registers), all years to 16 March 2015. We conducted complementary searches in PubMed and trials registries. Supplementary searches included reference lists of included studies, citation indexes, personal communication to the authors of all included studies and grey literature searches in OpenSIGLE. We applied no restrictions on date, language or publication status.

Selection criteria

We included all relevant randomised controlled trials (RCTs) focusing on adults with a formal diagnosis of panic disorder with or without agoraphobia. We considered the following psychological therapies: psychoeducation (PE), supportive psychotherapy (SP), physiological therapies (PT), behaviour therapy (BT), cognitive therapy (CT), cognitive behaviour therapy (CBT), third-wave CBT (3W) and psychodynamic therapies (PD). We included both individual and group formats. Therapies had to be administered face-to-face. The comparator interventions considered for this review were: no treatment (NT), wait list (WL) and attention/psychological placebo (APP). For this review we considered four short-term (ST) outcomes (ST-remission, ST-response, ST-dropouts, ST-improvement on a continuous scale) and one long-term (LT) outcome (LT-remission/response).

Data collection and analysis

As a first step, we conducted a systematic search of all relevant papers according to the inclusion criteria. For each outcome, we then constructed a treatment network in order to clarify the extent to which each type of therapy and each comparison had been investigated in the available literature. Then, for each available comparison, we conducted a random-effects meta-analysis. Subsequently, we performed a network meta-analysis in order to synthesise the available direct evidence with indirect evidence, and to obtain an overall effect size estimate for each possible pair of therapies in the network. Finally, we calculated a probabilistic ranking of the different psychological therapies and control conditions for each outcome.

Main results

We identified 1432 references; after screening, we included 60 studies in the final qualitative analyses. Among these, 54 (including 3021 patients) were also included in the quantitative analyses. With respect to the analyses for the first of our primary outcomes, (short-term remission), the most studied of the included psychological therapies was CBT (32 studies), followed by BT (12 studies), PT (10 studies), CT (three studies), SP (three studies) and PD (two studies).

The quality of the evidence for the entire network was found to be low for all outcomes. The quality of the evidence for CBT vs NT, CBT vs SP and CBT vs PD was low to very low, depending on the outcome. The majority of the included studies were at unclear risk of bias with regard to the randomisation process. We found almost half of the included studies to be at high risk of attrition bias and detection bias. We also found selective outcome reporting bias to be present and we strongly suspected publication bias. Finally, we found almost half of the included studies to be at high risk of researcher allegiance bias.

Overall the networks appeared to be well connected, but were generally underpowered to detect any important disagreement between direct and indirect evidence. The results showed the superiority of psychological therapies over the WL condition, although this finding was amplified by evident small study effects (SSE). The NMAs for ST-remission, ST-response and ST-improvement on a continuous scale showed well-replicated evidence in favour of CBT, as well as some sparse but relevant evidence in favour of PD and SP, over other therapies. In terms of ST-dropouts, PD and 3W showed better tolerability over other psychological therapies in the short term. In the long term, CBT and PD showed the highest level of remission/response, suggesting that the effects of these two treatments may be more stable with respect to other psychological therapies. However, all the mentioned differences among active treatments must be interpreted while taking into account that in most cases the effect sizes were small and/or results were imprecise.

Authors' conclusions

There is no high-quality, unequivocal evidence to support one psychological therapy over the others for the treatment of panic disorder with or without agoraphobia in adults. However, the results show that CBT - the most extensively studied among the included psychological therapies - was often superior to other therapies, although the effect size was small and the level of precision was often insufficient or clinically irrelevant. In the only two studies available that explored PD, this treatment showed promising results, although further research is needed in order to better explore the relative efficacy of PD with respect to CBT. Furthermore, PD appeared to be the best tolerated (in terms of ST-dropouts) among psychological treatments. Unexpectedly, we found some evidence in support of the possible viability of non-specific supportive psychotherapy for the treatment of panic disorder; however, the results concerning SP should be interpreted cautiously because of the sparsity of evidence regarding this treatment and, as in the case of PD, further research is needed to explore this issue. Behaviour therapy did not appear to be a valid alternative to CBT as a first-line treatment for patients with panic disorder with or without agoraphobia.

PLAIN LANGUAGE SUMMARY

Psychological therapies for the treatment of panic disorder with or without agoraphobia

Why is this review important?

Many people suffer from panic disorder. Panic disorder can occur on its own or with agoraphobia. People with panic disorder experience recurring panic attacks. During a panic attack people feel the sudden onset of intense fear alongside a series of bodily symptoms such as a racing heart, chest pain, sweating, shaking, dizziness, flushing, stomach churning, faintness and breathlessness. People with agoraphobia feel an intense fear of developing a panic attack in situations where escape might be difficult or embarrassing. This fear often leads to the avoidance of such situations.

There are many different types of talking therapies that are used to treat panic disorder with or without agoraphobia. However it is not clear whether certain talking therapies are more effective than others at treating panic disorder with or without agoraphobia. In this review we compared the effectiveness of different types of talking therapy.

Who will be interested in this review?

People with panic disorder with or without agoraphobia.

Friends and family of people with panic disorder with or without agoraphobia.

General practitioners, psychiatrists and psychologists.

Professionals working in adult mental health services.

What questions does this review aim to answer?

Are any of the included psychological therapies more effective and better tolerated than others in the rapid reduction of panic/agoraphobia symptoms?

Can any of the included psychological therapies guarantee better results one year after termination?

Which studies were included in the review?

We searched medical databases up to 16 March 2015 to find all studies (specifically randomised controlled trials) of talking therapies in the treatment of panic disorder with or without agoraphobia. To be included in the review studies had to include people with a clear diagnosis of panic disorder with or without agoraphobia.

We included 60 studies in the review. Fifty-four of the included studies (involving 3021 participants) were used in numerical analyses. The review authors rated the overall quality of the studies as low to very low.

What does the evidence from the review tell us?

The results of the review show that in general talking therapies are more effective than no treatment. There was no strong evidence to support one talking therapy over the others for the treatment of panic disorder with or without agoraphobia in adults. However, there was some low-quality evidence in favour of cognitive behaviour therapy (CBT), psychodynamic therapy and supportive psychotherapy over other talking therapies for short-term remission and short-term reduction in symptoms. The results concerning supportive psychotherapy should, however, be treated with caution because of the small amount of evidence available about this treatment. On the other hand, beyond the evidence regarding its efficacy, psychodynamic therapy also showed promising results in terms of tolerability: as a way of assessing how well people tolerated the talking therapies, we assessed short-term dropout rates. We found that there were fewer dropouts in psychodynamic therapy and third-wave CBT, suggesting that people tolerate these therapies better than other therapies.

What should happen next?

More high-quality research is needed to be able to fully compare the effectiveness of different talking therapies. In particular, more new studies are needed that compare the specific talking therapies CBT, psychodynamic therapy and supportive psychotherapy for the treatment of panic disorder with or without agoraphobia.

Summary of findings for the main comparison. Cognitive behaviour therapy compared to no treatment for panic disorder with or without agoraphobia in adults

Cognitive behaviour therapy compared to no treatment for panic disorder with or without agoraphobia in adults

Patient or population: adult patients with panic disorder with or without agoraphobia

Setting: outpatients
Intervention: cognitive behaviour therapy (CBT)
Comparison: no treatment (NT)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	No treatment (NT)	Cognitive-behavioral therapy (CBT)			
Short-term remission (follow-up: mean 3 months)	Study population 36 per 100	61 per 100 (24 to 89)	OR 2.78 (0.54 to 14.29)	No direct evidence available: indirect evidence based on 357 participants (2 studies via CT; 5 studies via PT)[a]	⊕⊖⊖⊖ VERY LOW [1,2,3]
Short-term response (follow-up: mean 3 months)	Study population 36 per 100	80 per 100 (42 to 97)	OR 7.14 (1.25 to 50)	No direct evidence available: indirect evidence based on 357 participants (2 studies via CT; 5 studies via PT)[a]	⊕⊖⊖⊖ VERY LOW [1,2,3]
Short-term dropouts (follow-up: mean 3 months)	Study population 1 per 100 (no events were observed in the available direct evidence: this percentage was assumed in order to calculate the corresponding risk)	6 per 100 (0 to 50)	OR 6.25 (0.26 to 100)	No direct evidence available: indirect evidence based on 278 participants (4 studies via PT)[a]	⊕⊖⊖⊖ VERY LOW [1,2,3]
Long-term remission/response (follow-up: mean 12 months)	*No data available for this comparison*	-	-	-	-

Short-term improvement as measured on a continuous scale (follow-up: mean 3 months)	The mean short-term improvement as measured on a continuous scale in the control group was 0	The mean short-term improvement, measured on a continuous scale as SMD (NMA results), was -0.83 (95% CI -1.5 to -0.16), indicating a large effect size in favour of CBT	-	27 participants (1 RCT)	⊕⊕⊖⊖ LOW [4,5]

Reported ORs are derived (as reciprocal values) from the results of network meta-analyses presented in Table 1, Table 4 and Table 7 (for ST-remission and ST-response we used the results of NMA adjusted for SSE). Reported SMD is derived from the results of network meta-analysis presented in Table 10.

[a]Indirect comparison is performed using all possible intermediate comparisons in the network. For brevity we report the number of studies contributing indirect evidence only via a single intermediate comparator.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **NMA:** network meta-analysis; **OR:** odds ratio; **RCT:** randomised controlled trial; **SMD:** standardised mean difference

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

We downgraded the quality of the evidence one step at a time when one or more of the following criticisms was present:
[1]Only indirect evidence is available for this outcome.
[2]Wide 95% CI.
[3]The risk of bias of indirect evidence is often unclear.
[4]The risk of bias in the included study was unclear in almost every domain.
[5]Only one study, with a small sample size, available for direct comparison.

Summary of findings 2. Cognitive behaviour therapy compared to supportive psychotherapy for panic disorder with or without agoraphobia in adults

Cognitive behaviour therapy compared to supportive psychotherapy for panic disorder with or without agoraphobia in adults

Patient or population: adult patients with panic disorder with or without agoraphobia

Setting: outpatients
Intervention: cognitive behaviour therapy (CBT)
Comparison: supportive psychotherapy (SP)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Supportive psychotherapy (SP)	Cognitive behaviour therapy (CBT)			
Short-term remission (follow-up: mean 3 months)	Study population 38 per 100	 29 per 100 (13 to 52)	OR 0.67 (0.25 to 1.82)	176 (3 RCTs)	⊕⊕⊖⊖ LOW [1,2]
Short-term response (follow-up: mean 3 months)	Study population 32 per 100	 34 per 100 (16 to 60)	OR 1.12 (0.4 to 3.26)	176 (3 RCTs)	⊕⊖⊖⊖ VERY LOW [1,2,3,4]
Short-term dropouts (follow-up: mean 3 months)	Study population 46 per 100	 35 per 100 (19 to 55)	OR 0.64 (0.28 to 1.43)	176 (3 RCTs)	⊕⊖⊖⊖ VERY LOW [1,2,3]
Long-term remission/response (follow-up: mean 12 months)	Study population 24 per 100	 40 per 100 (19 to 65)	OR 2.09 (0.73 to 5.98)	80 (1 RCT)	⊕⊕⊖⊖ LOW [1,5]
Short-term improvement as measured on a continuous scale (follow-up: mean 3 months)	The mean short-term improvement as measured on a continuous scale in the control group was 0	The mean short-term improvement, measured on a continuous scale as SMD (NMA results), was -0.05 (95% CI -0.56 to 0.47), indicating almost no difference between CBT and SP (the negative value of the point estimate indicates a slight trend in favour of CBT)	-	152 (3 RCTs)	⊕⊕⊖⊖ LOW [1,2]

Reported ORs and SMD are derived from the network meta-analyses.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CI: confidence interval; **NMA:** network meta-analysis; **OR:** odds ratio; **RCT:** randomised controlled trial; **SMD:** standardised mean difference

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect

Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

We downgraded the quality of the evidence one step at a time when one or more of the following criticisms was present:
[1]Risk of bias for the included studies was in many cases unclear.
[2]Only a few studies available for direct comparison. 95% CI still wide and non-significant even after combining direct and indirect evidence.
[3]Results were inconsistent across studies, although with wide confidence intervals.
[4]ST-response data were imputed from the continuous outcome for all the included studies.
[5]Only one study available for direct comparison. 95% CI still wide and non-significant even after combining direct and indirect evidence.

Summary of findings 3. Cognitive behaviour therapy compared to psychodynamic psychotherapy for panic disorder with or without agoraphobia in adults

Cognitive behaviour therapy compared to psychodynamic psychotherapy for panic disorder with or without agoraphobia in adults

Patient or population: adult patients with panic disorder with or without agoraphobia

Setting: outpatients
Intervention: cognitive behaviour therapy (CBT)
Comparison: psychodynamic psychotherapy (PD)

Outcomes	Illustrative comparative risks* (95% CI)		Relative effect (95% CI)	No of participants (studies)	Quality of the evidence (GRADE)
	Assumed risk	Corresponding risk			
	Psychodynamic psychotherapy (PD)	Cognitive behaviour therapy (CBT)			
Short-term remission (follow-up: mean 3 months)	Study population		OR 0.94 (0.27 to 3.45)[a]	54 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,2,3,4]
	44 per 100	43 per 100 (18 to 73)			
Short-term response (follow-up: mean 3 months)	Study population		OR 1.05 (0.28 to 4)[a]	54 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,2,3,4]
	47 per 100	48 per 100 (20 to 78)			
Short-term dropouts (follow-up: mean 3 months)	Study population		OR 1.92 (0.56 to 6.67)[a]	54 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,2,3,4]
	19 per 100	32 per 100			

		(12 to 62)			
Long-term remission/response (follow-up: mean 12 months)	Study population		OR 1.25 (0.37 to 4.17)[a]	54 (1 RCT)	⊕⊕⊖⊖ LOW [1,4]
	50 per 100	56 per 100 (27 to 81)			
Short-term improvement as measured on a continuous scale (follow-up: mean 3 months)	The mean short-term improvement as measured on a continuous scale in the control group was 0	The mean short-term improvement as measured on a continuous scale in the intervention group was 0.17 standard deviations higher (0.5 lower to 0.83 higher), indicating a small effect size in favour of PD[b]	-	54 (1 RCT)	⊕⊖⊖⊖ VERY LOW [1,3,4,5]

[a] Reported ORs are derived (as reciprocal values) from the results of network meta-analyses presented in Table 1, Table 4, Table 7 and Table 13 (for ST-remission and ST-response we used the results of NMA adjusted for SSE).

[b] Reported SMD is derived from the results of network meta-analysis presented in Table 10.

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
CBT: cognitive behaviour therapy; **CI:** confidence interval; **NMA:** network meta-analysis; **OR:** odds ratio; **PD:** psychodynamic psychotherapy; **PT:** physiological therapies; **RCT:** randomised controlled trial; **SMD:** standardised mean difference

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

We downgraded the quality of the evidence one step at a time when one or more of the following criticisms was present:
[1]The available direct evidence was affected by a high risk of bias in various important domains.
[2]Relevant (although non-significant) inconsistency was found in the loop PD-CBT-PT.
[3]Indirect evidence importantly influences the NMA results.
[4]Only one study available for direct comparison. 95% CI still wide and non-significant even after combining direct and indirect evidence.
[5]Statistically significant inconsistency was found in the loop PD-CBT-PT.

Summary of findings 4. Network meta-analysis rankings of psychological therapies for panic disorder with or without agoraphobia in adults

NMA Rankings of psychological therapies for panic disorder with or without agoraphobia in adults
Patient or population: adult patients with panic disorder with or without agoraphobia

Setting: outpatients
Intervention: psychoeducation (PE), supportive psychotherapy (SP), physiological therapies (PT), behaviour therapy (BT), cognitive therapy (CT), cognitive-behaviour therapy (CBT), third-wave CBT (3W), psychodynamic psychotherapy (PD)
Comparison: no treatment (NT), waiting list (WL), attention/psychological placebo (APP)

Outcomes	Treatment hierarchy (in descending order)	№ of participants (studies)	Quality of the evidence (GRADE)
Short-term remission (follow up: mean 3 months)	(SP)-CBT-PD-CT-BT-PT-NT-WL	2491 (40 RCTs)	⊕⊕⊖⊖ LOW [1,2]
Short-term response (follow up: mean 3 months)	CBT-PD-(SP)-BT-PT-WL-CT-NT	2240 (37 RCTs)	⊕⊕⊖⊖ LOW [1,2]
Short-term dropouts (follow up: mean 3 months)	NT-PD-WL-3W-CBT-APP-PE-PT-CT-BT-SP	2535 (47 RCTs)	⊕⊕⊖⊖ LOW [1,3]
Long-term remission/response (follow up: mean 12 months)	CBT-PD-PT-BT-SP-CT	464 (9 RCTs)	⊕⊕⊖⊖ LOW [1,3]
Short-term improvement as measured on a continuous scale (follow up: mean 3 months)	(PD)-CBT-SP-CT-3W-BT-PT-NT-WL	2318 (57 RCTs)	⊕⊕⊖⊖ LOW [1,3]

Reported rankings are based on absolute SUCRA values, which are derived from network meta-analyses (NMA).

The ranking of treatments reported in parenthesis must be interpreted with caution, because the evidence supporting those rankings is either too scarce or hampered by relevant inconsistency.

The assessment of quality of evidence has been made by adapting the GRADE tool, designed for pairwise meta-analyses, to network meta-analyses, as suggested in Salanti 2014.

GRADE Working Group grades of evidence
High quality: We are very confident that the true effect lies close to that of the estimate of the effect
Moderate quality: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low quality: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect
Very low quality: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect

[1] Downgraded for study limitations because the risk of bias was unclear or high in more than one important domain for many of the included studies.
[2] Downgraded for imprecision because too few comparisons remained clinically important after adjusting the results of NMA for SSE (See additional Table 1 and Table 4)
[3] Downgraded for imprecision because too few comparisons showed clinically important results in NMA (See additional Table 7, Table 10 and Table 14)

BACKGROUND

Description of the condition

A panic attack is a discrete period of fear or anxiety that has a rapid onset, reaches a peak within 10 minutes and in which at least four of 13 characteristic symptoms are experienced. Many of these symptoms involve bodily systems, such as racing heart, chest pain, sweating, shaking, dizziness, flushing, stomach churning, faintness and breathlessness. Further recognised panic attack symptoms involve fearful cognitions, such as the fear of collapse, going mad or dying, and derealisation (APA 2000).

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) (APA 2000), panic disorder is characterised by the presence of recurrent unexpected panic attacks, of which at least one has been followed by one month (or more) of persistent concern about having additional attacks, worry about the implications of the attack (or its consequences) or a significant change in behaviour related to the attacks.

Panic disorder is common in the general population, with a lifetime prevalence of 1% to 4% (Bijl 1998; Eaton 1994). In primary care settings panic syndromes have been reported to have a prevalence of around 10% (King 2008). The aetiology of panic disorder is not fully understood and is probably heterogeneous. Biological theories incorporate the faulty triggering of an inbuilt anxiety response, possibly a suffocation alarm. Evidence for this comes from biological challenge tests (lactate and carbon dioxide trigger panic in those with the disorder) and from neuroimaging studies that show activation of fear circuits, such as involving the periaqueductal grey matter (Gorman 2000).

Agoraphobia is anxiety about being in places or situations from which escape might be difficult or embarrassing, or in which help may not be available in the event of having a panic attack (APA 2000). Agoraphobia can occur with panic disorder and in the general population about one quarter of people suffering from panic disorder also have agoraphobia but this proportion is much higher in the clinical samples (Kessler 2006). The presence of agoraphobia is associated with increased severity and worse outcomes (Kessler 2006). There are several risk factors that predict the development of agoraphobia in people suffering from panic disorder including female gender, more severe dizziness during panic attacks, cognitive factors, dependent personality traits and social anxiety disorder (Starcevic 2009).

Panic disorder is more common among women, with a 2:1 ratio; in the case of panic disorder with agoraphobia the ratio rises to 3:1. Most typically, the disorder strikes between late adolescence and 35 years of age; early or late onset is possible, although less common (APA 2000).

Panic disorder, with or without agoraphobia, is highly comorbid with other psychiatric disorders such as drug dependence, major depression, bipolar I disorder, social phobia, specific phobia and generalised anxiety disorder (Grant 2006). It is estimated that generalised anxiety disorder co-occurs in 68% of people with panic disorder, whilst major depression has a prevalence of 24% to 88% among people with panic disorder (Starcevic 2009).

Description of the intervention

Recent guidelines from the National Institute for Health and Clinical Excellence recommend three types of intervention in the care of individuals with panic disorder, any of which should be offered promptly, taking into account the preference of the patient (NICE 2011). According to the NICE guidelines, the interventions that have evidence for the longest duration of effect, in descending order, are psychological therapy, pharmacological therapy (antidepressant medication) and self help.

A psychological therapy can be defined as a therapeutic interaction between a trained professional and a patient (or a group of patients) by way of their verbal and non-verbal communication for the purpose of ameliorating the sufferings on the part of the patient(s).

Although NICE guidelines recommend the use of cognitive behaviour therapy (CBT) for the treatment of panic disorder, many other psychological therapies have been proposed as viable therapeutic options. Each therapy is characterised by a certain theoretical framework, according to which a set of therapeutic ingredients (or 'components') and technical features can be defined and briefly described as follows.

- Psychoeducation consists of providing patients with information about their psychological disease. In this context, it can be explained to patients that their symptoms can be interpreted in the light of a certain cause-effect model, according to a more general theoretical framework that can vary across the different psychological approaches.
- Supportive psychotherapy is a dyadic treatment that uses direct measures to ameliorate symptoms and maintain, restore or improve self esteem, ego function and adaptive skills (according to the manual of Winston, Rosenthal and Pinsker (Winston 2004)). Although different techniques can be used (e.g. encouragement, rationalising and reframing, anticipatory guidance, etc.) therapeutic alliance represents the most important element of the therapy (Winston 2004). Rogerian client-centred psychotherapy is probably the most representative example of supportive psychotherapy (Rogers 1980). In this approach, within the context of a warm, empathic and non-directive therapeutic relationship clients are led to become aware of their true feelings and to fully accept themselves as they are, including imperfections and dysfunctions.
- Physiological therapies are represented by a set of different possible treatments that use some kind of physical training (e.g. breathing retraining, relaxation techniques, biofeedback) in order to help the patient to control the physiological manifestations of anxiety. Among the physiological therapies proposed for the treatment of panic disorder, breathing retraining and relaxation techniques are probably the most studied. Within the context of breathing retraining, different strategies have been proposed, although most manuals and studies describe instructions in abdominal breathing as their central technique (Meuret 2012). Progressive muscle relaxation (PMR), as formalised by Bernstein and Borkovec (Bernstein 1973), can be taught to panic patients in order to reduce general tension and achieve a body state that lowers the risk for stressors to elicit panic. The so-called applied relaxation

is a slightly different form of physiological therapy in which relaxation training and exposure are combined (Ost 1987).

- The behavioural therapy of panic disorder consists of graded exposure to the body sensations that accompany panic ('interoceptive exposure') or to situations perceived as threatening ('in vivo exposure', 'imagery exposure', 'virtual reality exposure'), or both, in order to progressively reduce the patient's apprehensive reaction towards them.
- Cognitive therapy finds its roots in the work of Albert Ellis and Aaron Beck. Its main component is represented by cognitive restructuring, a psychotherapeutic process of learning to identify and modify irrational or maladaptive thoughts using strategies such as Socratic questioning, thought recording and guided imagery.
- CBT for panic disorder is usually administered according to the manuals of Clark 1986b and Barlow 2000b. Its main components are represented by psychoeducation, breathing retraining, PMR, cognitive restructuring, behavioural experiments, interoceptive exposure and in vivo exposure.
- The so-called 'third-wave' therapies are represented by a set of different therapies (e.g. mindfulness-based cognitive therapy, acceptance and commitment therapy, compassionate mind training, extended behavioural activation, metacognitive therapy, schema therapy), all originating from the cognitive behavioural approach but compared to which more importance is given to the form, rather than the content, of patients' thoughts. By focusing on the function of cognition, third-wave therapies aim to help patients to develop more adaptive emotional responses to situations. When mindfulness and acceptance are applied to anxiety disorders, the aim is for the individual to be able to observe symptomatic processes without overly identifying with them or without reacting to them in ways that cause further distress (Roemer 2008). A systematic review and meta-analysis of mindfulness-based and acceptance-based interventions for anxiety disorders has been recently published (Vøllestad 2012).
- Psychodynamic therapies consist of a set of psychological therapies, different in length and depth, represented by psychoanalysis (as conceptualised by S Freud) and its further developments. According to psychodynamic psychotherapy, psychological symptoms can be seen as manifestations of intrapsychic or unconscious conflicts; these therapies use different therapeutic strategies (e.g. unconscious contents exploration, dream analysis, analysis of past experiences, analysis of parental relationships, analysis of transference, analysis of resistances) in order to reveal, interpret and resolve such conflicts. A brief panic-focused psychodynamic psychotherapy is described in Milrod 1997. A sightly different approach, derived from psychodynamic theories, is represented by so-called emotion-focused therapy (EFT) in which the therapist is viewed as an 'emotion coach' who works to enhance emotion-focused coping by helping people become aware of, accept and make sense of their emotional experience (Greenberg 2004).

A further level of distinction among psychological therapies concerns the form of delivery of the intervention. In this regard, NICE guidelines suggest that the intervention (CBT) should be optimally delivered in the form of one- to two-hour weekly sessions, for a total of seven to 14 hours, within a maximum of four months since commencement. However, different variables have been and still remain the subject of investigation with regard to cost-effectiveness analyses. A number of variables including the number of sessions, the duration of treatment and the therapeutic setting (group versus individual; face-to-face versus remote versus self help) have been explored.

How the intervention might work

The main features and rationale of the psychological therapies considered for this review can be summarised as follows.

The rationale of psychoeducation is that providing anxious patients with a better understanding of their sufferings may in itself lead to symptom relief (Clark 1985; Sorby 1991). This may be especially important in panic disorder, where the cognitive coping mechanisms of the patients are disrupted and where anticipatory anxiety may cause additional attacks (Dannon 2002). In this sense, as the authors further suggest, a psychoeducational intervention may increase the patients' sense of control leading to a reduction of catastrophic thoughts and emotions.

Supportive psychotherapy is non-specific in nature, so it is not designed for the treatment of a psychiatric disorder in particular. In this sense, the supportive treatment of panic disorder and agoraphobia does not differ from the treatment of any other disorder. Although scarce, the available body of evidence does not exclude a possible role of supportive psychotherapy in the treatment of agoraphobia (Klein 1983; Zitrin 1978); its efficacy in the treatment of panic disorder without agoraphobia still remains unclear.

Respiratory abnormalities, with particular regard to hyperventilation and hypocapnia, have been postulated as being important factors in the development or maintenance of panic disorder (Klein 1993; Ley 1985; for a review see Meuret 2010b). According to the model proposed by Ley 1985, panic attacks are caused by acute states of hypocapnia in a positive feedback loop between hyperventilation and anxiety. Therefore, amelioration of panic symptoms is expected when patients achieve reductions in transient and sustained hypocapnia. Results on the efficacy of breathing training in the treatment of panic disorder are mixed (Meuret 2010b). The purpose of applied relaxation is to teach the patient to observe the very first signs of a panic attack and to apply a rapid and effective relaxation technique to cope with, and eventually abort, these symptoms before they develop into a panic attack. In a direct comparison with regard to remission from panic disorder at the end of treatment, applied relaxation was not found to be significantly better than PMR, although it performed better than PMR on six out of 11 measures (Ost 1988).

Behaviour therapy (BT) is characterised by the use of some kind of exposure in order to modify dysfunctional behaviours that may contribute to the development and persistence of psychological symptoms. The principle of exposure in the treatment of phobic disorders is to persuade the patient to enter and stay in his or her phobic situation until he or she feels better, and to do this repeatedly until it becomes so customary that the situation no longer holds terror (Marks 1981). There is evidence that exposure strategies alone are effective in the treatment of panic disorder (Gloster 2011; Ost 2004; Williams 1996).

In the case of cognitive therapy (CT) for panic disorder, it has been proposed that panic attacks result from the catastrophic

misinterpretation of certain bodily sensations (Clark 1986a). This involves the sufferer perceiving sensations involved in normal anxiety responses as much more dangerous than they really are, for example perceiving palpitations as evidence of impending heart attack. The cognitive approach would involve identifying patients' negative interpretations of the bodily sensations experienced in panic attacks, suggesting alternative non-catastrophic interpretations of the sensations, and then helping the patient to test the validity of these alternative interpretations. As pointed out in a recent review (Meuret 2012), CT is often intermingled with behavioural techniques (for example, 'behavioural experiments', 'hypothesis testing', 'instructions' involving exposure), which complicates the testing of the efficacy of CT in its 'pure' form. Nonetheless, there is some evidence that training in cognitive procedures in full isolation from exposure and behavioural procedures is efficacious in reducing aspects of panic (Beck 1994; Meuret 2010a; Salkovskis 1991; Van den Hout 1994).

Cognitive behaviour therapy (CBT) combines elements of both in order to reduce emotional distress and psychological symptoms, assuming that cognitions, behaviours and emotions are interrelated. A fairly consistent body of evidence exists in support of the efficacy of CBT for panic disorder, administered either in individual or group sessions (among others: Clark 1999; Dow 2000; Hendriks 2010; Telch 1993). Furthermore, a growing body of evidence supports the efficacy of self administered versions (for example, book-based, internet-based) of this psychological therapy (Carlbring 2006; Nordin 2010; Wims 2010).

As summarised in Ludwig 2008, mindfulness involves attending to relevant aspects of experience in a non-judgemental manner. The goal of mindfulness is to maintain awareness moment by moment, disengaging oneself from strong attachment to beliefs, thoughts or emotions, thereby developing a greater sense of emotional balance and well-being. An aim of mindfulness practice is to take greater responsibility for one's life choices. Although scarce, some evidence exists in support of the efficacy of this therapy for the treatment of generalised anxiety disorder and panic disorder (Kim 2009; Lee 2007). As originally developed (Hayes 1999), acceptance and commitment therapy (ACT) was intended for the treatment of psychopathology in general rather than a specific disorder in particular. ACT conceptualises psychological events as a set of ongoing interactions between whole organisms and historically and situationally defined contexts. Removal of a client's problematic behaviours from the contexts that participate in that event (for example, merely analysing manifested behavioural symptoms themselves) is thought to miss the nature of the problem and the avenues for its solution. ACT clients are therefore encouraged to embrace a passionate and ongoing interest in how to live according to their values. In ACT there is a conscious posture of openness and acceptance toward all psychological events, even if they are formally 'negative', 'irrational' or even 'psychotic'. For example, if the client feels trapped, frustrated, confused, afraid, angry or anxious, the ACT stance suggests this is not so much a problem as it is an opportunity to work on how powerful events in the here and now can become barriers to growth (Hayes 2004). Some evidence supports ACT possibly being as effective as CBT in the treatment of anxiety disorders including panic disorder (Arch 2012).

Following a psychodynamic approach, Busch and colleagues proposed that during childhood, a sense of fearful dependency on the parent may lead to the development of anger towards him or her (Busch 1996). As a consequence, a vicious cycle develops in which the child's anger threatens the needed tie to the parent and thereby increases fearful dependency, which promotes further frustration and rage at the parent. This cycle may then recur in adulthood when threats to attachment trigger intense feelings of abandonment, anger and anxiety, leading to the development of the disorder. The aim of psychodynamic psychotherapy is to address such underlying psychological factors in order to obtain an improvement of panic symptoms. Although only a few studies have explored the effects of psychodynamic psychotherapy for panic disorder, the available evidence suggests the viability of this approach as a valid therapeutic option (Milrod 2007; Wiborg 1996).

Why it is important to do this review

A previous Cochrane meta-analysis comparing combined psychological therapy plus antidepressants versus psychological therapy alone or pharmacotherapy alone showed the superiority of combined therapy over either monotherapies in the short term, and of combined therapy and psychological therapy alone over pharmacotherapy alone in the long term, thus suggesting that either combined therapy or psychological therapy alone can be chosen as first-line treatment for panic disorder with or without agoraphobia (Furukawa 2007). In particular, behavioural and cognitive behavioural psychological therapies showed the strongest evidence. Another meta-analysis, aimed at analysing the efficacy of psychological interventions versus control conditions in the treatment of panic disorder with or without agoraphobia (Sánchez-Meca 2010), showed a general efficacy of psychological therapies over different clusters of symptoms, with the most consistent results in favour of the combination of exposure strategies with relaxation training or breathing retraining techniques, or both. The study conducted by Sánchez-Meca et al revealed the presence of substantial heterogeneity among included studies (I^2 = 70.4%). Exploratory secondary analyses suggested that variables such as type of therapy and type of control group may explain part of the observed heterogeneity. The observed degree of heterogeneity due to differences in the psychological therapies suggests that some psychological therapies may be more effective than others in the treatment of the disorder. However, both the existence and the eventual magnitude of such differences remain unclear. This is partly due to the presence of methodological diversity among available studies; as suggested by Sánchez-Meca et al, the type of control group may significantly influence the measured effect size, limiting the possibility of drawing conclusions. A further consideration is that only a few trials compared different psychological approaches with each other and, more generally, psychological therapies have not been all equally investigated.

In an attempt to overcome these issues, in this review we performed a network meta-analysis (NMA), also known as multiple treatment meta-analysis, in which eight different forms of psychological therapy and three forms of a control condition (see Types of interventions) have been independently compared with each other. We expected this methodological strategy to reduce the amount of heterogeneity that was observed in previous studies. Furthermore, by synthesising the available direct and indirect evidence via NMA, it was possible to obtain an overall effect size estimate for each

possible pair of therapies in the network, even for interventions that had not been directly compared with each other in previous trials. Finally, it was possible to calculate a probabilistic ranking in order to help the identification of those interventions that are more likely to be more effective than others in the treatment of panic disorder.

This review along with several others in progress, contributes to the production of a comprehensive portfolio of Cochrane reviews in the area of panic disorder (Furukawa 2007; Guaiana 2013; Guaiana 2013a; Guaiana 2013b; Watanabe 2009; Xiao 2011).

OBJECTIVES

To assess the comparative efficacy and acceptability of different psychological therapies and different control conditions for panic disorder, with or without agoraphobia, in adults.

METHODS

Criteria for considering studies for this review

Types of studies

All relevant randomised controlled trials (RCTs).

We included cluster-randomised trials when the effects of clustering were taken account of (however, we found no such cases).

We included cross-over randomised trials, but we only considered results from the first randomisation period.

We included studies in which the replacement of dropouts was allowed as long as replacements were low in number (less than 15% of the final sample) and evenly distributed among treatment arms.

We excluded quasi-randomised controlled trials (in which treatment assignment was decided through methods such as alternate days of the week).

Types of participants

Age range

Patients, aged 18 years or older, of both sexes. We included studies that included some participants under the age of 18 as long as at least 80% of patients were aged 18 years or above.

Diagnosis

We included studies that had enlisted participants with a primary diagnosis of panic disorder with or without agoraphobia diagnosed according to any of the following criteria: Feighner criteria (Feighner 1972), Research Diagnostic Criteria (Spitzer 1978), DSM-III (APA 1980), DSM-III-R (APA 1987), DSM-IV (APA 2000), DSM-5 (APA 2013) or ICD-10 (WHO 1992). When ICD-10 or DSM-5 were used, in which panic disorder and agoraphobia are separately diagnosable, this review focused on panic disorder comorbid with or without agoraphobia. We took the latter decision in order to be concordant with the current body of literature, most of which used DSM-III-R or DSM-IV and little, if any, used ICD-10 or DSM-5.

There is evidence that over 95% of patients with agoraphobia who are seen clinically suffer from panic disorder as well (Goisman 1995). According to this finding, we included studies focusing on agoraphobia, rather than panic disorder, if operationally diagnosed according to the above-mentioned criteria and when it could be safely assumed that at least 80% of the participants were suffering from panic disorder. We explored the effect of the inclusion of trials with different percentages of patients suffering from agoraphobia in a meta-regression analysis.

Setting

Participants must have been outpatients at the time of enrolment.

Previous treatment

We included both treatment-naive patients and patients who had already undergone some previous treatment (either psychological or pharmacological), as long as they satisfied the above-mentioned inclusion criteria. However, we excluded studies where all participants had shown resistance to previously administered psychological therapies.

Comorbidities

We included studies where participants had other anxiety disorders (for example, generalised anxiety disorder, specific phobias) or with subthreshold panic disorder if: 1) separate results for patients with panic disorder were reported and 2) randomisation was stratified by specific diagnoses. Stratification by diagnosis was not required if the total sample included at least 40 participants with panic disorder.

We included studies in which the participants had physical comorbidities. However, we excluded studies explicitly focusing on panic disorder or agoraphobia among patients with a certain physical comorbidity.

We excluded studies in which all participants had a concurrent primary diagnosis of Axis I or II disorders other than panic disorder or agoraphobia.

Types of interventions

For this review, we chose to focus on most representative psychological therapy schools (that is CBT and its components or developments, psychodynamic psychotherapy and supportive psychotherapy) and their control conditions.

Experimental interventions

We included the following psychological therapies.

1. PE: psychoeducation, intended as sessions in which patients were only provided information about their disease.
2. SP: supportive psychotherapy, with or without a psychoeducational component, intended as sessions in which patients were administered an active, although non-specific, psychological treatment.
3. PT: physiological therapies that used some kind of physical training (e.g. breathing retraining, progressive muscle relaxation, applied relaxation) in order to reduce the physiological manifestations of anxiety.
4. BT: behaviour therapy, with or without physiological components, aiming at patients' habituation to anxiety-provoking situations and sensations through some kind of exposure (e.g. interoceptive, in vivo).

5. CT: cognitive therapy, with or without physiological components and behavioural experiments, aiming at the modification of maladaptive thoughts through some kind of cognitive restructuring.
6. CBT: cognitive behaviour therapy, with or without physiological components, containing both cognitive and behavioural therapy elements.
7. 3W: third-wave CBT, including acceptance and commitment therapy, mindfulness-based therapy, and other so-called 'third-wave' therapies administered with or without other CBT components (e.g. exposure, cognitive restructuring, breathing retraining, muscle relaxation).
8. PD: psychodynamic therapies focused on revealing and resolving intrapsychic or unconscious conflicts.

When psychoeducation or psychological support, or both, were accompanied by any other psychological intervention, we classified the study arms according to the latter and we regarded psychoeducation and psychological support as components of that intervention.

Therapies could be of any length so that we accepted those given in a single session.

We included both individual and group therapies.

We included the so-called component studies (for example, dismantling studies) as long as each arm could be regarded as any of the above-defined experimental interventions compared against another experimental or comparator treatment. Eventually, study arms could be regarded as giving information about the same experimental intervention and thus be combined.

Therapies had to be administered face-to-face. We excluded therapies administered in their self help (for example, book, computer, Internet) or remote (for example, telephone, video-conference) versions. In the case of psychoeducation, the simple provision of informational material without any face-to-face session was not considered an active intervention but rather a comparator intervention, such as no psychological treatment or wait list (however, we found no similar cases).

We excluded combination therapies. However, we included studies in which a pharmacological co-administration was allowed as long as there were no systematic differences in drug administration between the study arms. The percentage of studies in which a drug co-administration was allowed, the percentage of studies that required a stabilisation of therapy and, in this latter case, the time required for stabilisation, is reported.

We excluded any other psychological approach (such as interpersonal therapy (IPT), eye movement desensitisation and reprocessing (EMDR) and Morita therapy) on the grounds that they do not meet the criteria for a CBT (and its components and developments), psychodynamic psychotherapy or supportive psychotherapy.

We excluded family therapy, couple therapy and other psychosocial interventions whose focus was not the individual but rather the family system or couple as a whole.

Comparator interventions

1. NT: no psychological treatment (participants received assessment only, with or without simple provision of informational material or minimal therapist contact, or both, and they knew that they would have not received the active treatment in question after the trial).
2. WL: wait list (participants received assessment, with or without simple provision of informational material or minimal therapist contact, or both, and they knew that they would have received the active treatment in question after the waiting phase).
3. APP: attention or psychological placebo (participants received a face-to-face inactive intervention*).

Given the general inconsistency of the definitions of comparator interventions among different studies, the attribution of a control group to one of these prespecified categories relied on its detailed description rather than on the name given by the authors. However, where a sufficiently detailed description was unavailable, either from the paper or by contacting the original authors, the attribution relied solely on the given definition. Particular inconsistency exists in the definition of what is intended for treatment as usual (TAU). When TAU was intended as no treatment, wait list or supportive psychotherapy, we classified groups accordingly.

*Attention placebo is defined as any form of inactive intervention designed by the original authors to be perceived as ineffective by patients; psychological placebo is defined as any form of inactive intervention designed by the original authors to be perceived as effective by patients. The inclusion of an intervention among attention or psychological placebo groups required the intervention to be inactive. Any form of active intervention was therefore included among experimental interventions even if defined as a control condition by the original authors.

We excluded studies in which a pharmacological placebo was either co-administered or used as the control condition.

In total we expected the network to have 11 nodes, each one representing an intervention or control (see Data synthesis).

Types of outcome measures

Primary outcomes

1. Short-term[a] remission[b] of panic disorder with or without agoraphobia
2. Short-term response[c] of panic disorder with or without agoraphobia
3. Dropouts for any reason in the short term (as a proxy for treatment acceptability)

(a) Short-term, i.e. within six months from treatment commencement. When multiple time point measures in the short term were available, we gave preference to measures at approximately three months after treatment commencement.

(b) 'Remission' was intended as a dichotomous outcome expressing the number of patients who reached a satisfactory end state as defined by global judgement by the original investigators. Examples are 'panic-free' and 'no or minimal symptom' according to the Clinical Global Impression Severity Scale (Guy 1976).

[c] 'Response' was intended as a dichotomous outcome expressing the number of patients who had a substantial improvement from baseline as defined by the original investigators. Examples are 'very much or much improved' according to the Clinical Global Impression (CGI) Change Scale (Guy 1976), more than 40% reduction in the score of the Panic Disorder Severity Scale (PDSS) (Shear 1997), and more than 50% reduction in the Fear Questionnaire Agoraphobia Subscale (FQ-ag) (Marks 1979).

When more than one index of remission or response was reported, we gave preference to the most global measure (e.g. in the case of remission, 'high end-state functioning' status was usually a more global index than 'panic-free' status); when more than one index was available but measures were equally 'global', we gave preference according to the same criteria used for the continuous scale outcome (see below). The actual measure entered into the meta-analysis is indicated in the table of included studies.

Secondary outcomes

4. Short-term improvement of panic disorder with or without agoraphobia as measured on a continuous scale[d]

5. Long-term[e] remission or response[f] of panic disorder with or without agoraphobia

[d] Examples are Panic Disorder Severity Scale (total score 0 to 28), Panic and Agoraphobia Scale (total score 0 to 45), Clinical Global Impression Severity Scale (1 to 7), Clinical Global Impression Change Scale (1 to 7), etc. When more than one scale was available in the paper, we gave preference in the following order:

- PDSS > Panic and Agoraphobia Scale (PAS) > ASI-R > ASI > ACQ > BSQ > other scales specific for panic disorder;
- CGI-S > CGI-I > GAS > GAF > other global scales;
- FQ-ag > FQ-global > Mobile Inventory for Agoraphobia-Avoidance-Alone (MI-AAL) > MI-Avoidance-Accompanied (MI-AAC) > other scales specific for agoraphobia only;
- panic frequency > panic severity > other scales specific for panic attacks only.

Once the scale was chosen, if both self and observer-rated assessments were available, we gave preference to the latter. The actual measure entered into the meta-analysis is indicated in the table of Characteristics of included studies.

[e] Long-term, i.e. six months or longer after treatment commencement, either on treatment discontinuation or on continued treatment (in the case of long-term therapies). When multiple time point measures in the long term were available, we gave preference to measures at approximately 12 to 15 months after treatment commencement. In the case of missing data at the long-term assessment, we considered studies for the analyses as long as dropouts were low in number (< 30% of the original sample) and evenly distributed across treatment arms.

[f] 'Response' and 'Remission' were intended as above. When both remission and response rates were reported, we considered the former. However, if remission rates were not reported but response rates were available, we used these for the analyses.

Search methods for identification of studies

CCDAN Specialised Register (CCDANCTR)

The Cochrane Depression, Anxiety and Neurosis Group (CCDAN) maintain two clinical trials registers at their editorial base in Bristol, UK: a references register and a studies-based register. The CCDANCTR-References Register contains over 40,000 reports of trials in depression, anxiety and neurosis. Approximately 50% of these references have been tagged to individual, coded trials. The coded trials are held in the CCDANCTR-Studies Register and records are linked between the two registers through the use of unique Study ID tags. Coding of trials is based on the EU-Psi coding manual. Please contact the CCDAN Trials Search Co-ordinator for further details. Reports of trials for inclusion in the Group's registers are collated from routine (weekly), generic searches of MEDLINE (1950-), EMBASE (1974-) and PsycINFO (1967-); quarterly searches of the Cochrane Central Register of Controlled Trials (CENTRAL) and review-specific searches of additional databases. Reports of trials are also sought from international trials registers c/o World Health Organization (WHO) trials portal (ICTRP), drug companies, handsearching of key journals, conference proceedings and other (non-Cochrane) systematic reviews and meta-analyses.

Details of CCDAN's generic search strategies can be found on the Group's website.

Electronic searches

We conducted the following searches (all years) to 16 March 2015.

We searched the CCDANCTR-Studies Register using the following terms:

Condition/Comorbidity = panic
AND
Intervention = (attention* or behav* or biblio* or biofeedback or cognitive or collaborative or contact or counsel* or desensiti* or educat* or expos* or feedback or ”group” or imag* or interpersonal or intervention or management or panic or prevention or psycho* or relaxation or self* or stress* or support* or *therap* or *train* or treatment or unclear or ”not stated”)

We searched the CCDANCTR-References Register using a more sensitive set of terms to identify additional untagged or uncoded reports of RCTs (Appendix 1).

We conducted a further search of the CCDANCTR to identify reports of studies for ‘Anxiety Disorders Not Otherwise Specified’ (ADNOS):

The CCDANCTR-Studies Register was searched for CONDITION = "Anxiety Disorder*"

We manually screened out pharma studies and studies in children and adolescents.

We searched the CCDANCTR-References Register using the following terms to identify additional untagged or uncoded reports of RCTs for ADNOS:

(“anxiety disorder*” and not (agoraphobi* or panic or (social and (anxi* or phobi*)) or generalised or generalized or obsessive or compulsive or OCD or PTSD or post-trauma* or “post trauma*” or posttrauma*)) +(terms for psychotherapies as listed in Appendix 1).

We manually screened out pharma studies and studies in children and adolescents retrieved from this sensitive search of the references register.

Supplementary searches

We conducted complementary searches in PubMed (Appendix 2) as well as in trials registries such as the WHO International Clinical Trials Registry Platform (http://apps.who.int/trialsearch/) and ClinicalTrials.gov (http://clinicaltrials.gov/).

There were no restrictions on date, language or publication status applied to the searches.

Searching other resources

Reference lists

We checked the reference lists of all included studies and relevant systematic reviews to identify additional studies missed from the original electronic searches.

Citation indexes

We conducted a citation search on the Web of Science to identify articles citing any of the included studies.

Personal communication

We contacted trialists and subject experts for information on unpublished or ongoing studies or to request additional trial data.

Grey literature

We searched the database OpenSIGLE (http://www.opengrey.eu/) to identify reports of trials not formally published in books or journals.

Data collection and analysis

Selection of studies

At least two out of three review authors (AP, AT, HI) examined the titles and abstracts of references identified by the electronic search strategies described above to check whether the study was likely to be relevant. We then obtained each potentially relevant study located in the search as a full article and the same two review authors independently assessed each for inclusion. In the case of discordance, we sought resolution by discussion. When disagreement could not be solved by discussion, arbitration was provided by a fourth author (TAF). Agreement between review authors in the study selection is reported. We evaluated the discordance in the selection of studies by quantifying both the percentage of agreement and Cohen's Kappa (k) (Cohen 1960). Where it was not possible to evaluate the study because of missing information, we classified the study as a 'Study awaiting assessment'. The reasons for the exclusion of trials are reported in the Characteristics of excluded studies table. Decisions made in the study selection process (along with number of references and studies, and reasons for exclusion of studies) are presented in a PRISMA flow diagram.

Data extraction and management

At least two out of three review authors (AP, AT, HI) used a structured, pilot-tested, Excel data collection form to independently extract the data from the included studies. Extracted data concerned: study design, administered interventions (format and timing of psychological therapy and control condition, therapist training, intervention components), participants' characteristics (diagnostic criteria, percentage of agoraphobic patients), outcomes, risk of bias and publication. Again, we resolved any disagreement either by discussion or by consultation of a fourth member of the review team (TAF). If necessary, we contacted authors of studies to obtain further clarification. Agreement between the data extractors with regard to primary outcomes is reported.

Assessment of risk of bias in included studies

At least two out of three review authors (AP, AT, HI) independently assessed the risk of bias of the included studies using the tool described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011). We assessed the following domains.

1. Random sequence generation and allocation concealment (selection bias).
2. Therapist and researcher allegiance, treatment fidelity (performance bias).
3. Blinding of outcome assessor (detection bias).
4. Incomplete outcome data reporting (attrition bias).
5. Selective outcome reporting (reporting bias).

We assessed and categorised the risk of bias, in each domain and overall, into:

- low risk of bias, plausible bias unlikely to seriously alter the results;
- high risk of bias, plausible bias that seriously weakens confidence in the results;
- unclear risk of bias, plausible bias that raises some doubt about the results.

Where inadequate details of randomisation and other characteristics of trials were provided, we classified the risk of bias as unclear, unless further information could be obtained by contacting the authors. If the assessors disagreed, we made the final rating by discussion or with the involvement of another member of the review group (TAF), if necessary. Agreement between the two independent raters in the 'Risk of bias' assessment is reported (see Risk of bias in included studies).

We assessed therapist and researcher allegiance, as well as treatment fidelity, as possible sources of performance bias. Blinding of therapists, the common way to minimise the risk of performance bias, is not feasible in these kinds of studies.

We evaluated the risk of detection bias for the first of the primary outcomes only. We classified studies as having a low risk of detection bias when the identification of a patient as a 'remitter' required at least one observer rating and the observer was blind to the treatment allocation.

We separately calculated risk of attrition bias for short-term and long-term outcomes, whenever such outcomes had been extracted. We classified a study as being at low risk of attrition bias when data for all randomised patients were available at short and long-term assessment. In the case of dropouts, a study may still be assessed as being at low risk of attrition bias when:

- missing outcome data were few and balanced in numbers across intervention groups, with similar reasons for missing data across groups;
- reasons for missing outcome data were unlikely to be related to true outcome;
- missing data had been imputed using appropriate methods (last observation carried forward (LOCF) was not considered an appropriate method in itself. It was considered appropriate only when the LOCF cases were few and balanced between arms).

Whenever possible, we retrieved study protocols in order to assess the risk of reporting bias. We considered a study to be at low risk of reporting bias when the study protocol was available and all of the study's prespecified (primary and secondary) outcomes that are of interest in the review had been reported in the prespecified way. When the study protocol was not available, we classified the study as being at unclear risk of reporting bias unless the reported information was enough to make a judgement (text of this nature was uncommon).

Measures of treatment effect

Dichotomous data

As the measure of treatment effect for binary outcomes we used the odds ratio (OR) and its 95% confidence interval (CI).

Continuous data

Since different studies have used different panic rating scales, we used the standardised mean difference (SMD) and its 95% confidence interval (CI).

Endpoint versus change data

We first planned to use scale endpoint data, which typically cannot have negative values and are easier to interpret from a clinical point of view. However, as a post hoc decision, we decided to use change data in an attempt to reduce the amount of heterogeneity due to the baseline imbalance found across studies. This decision, which is in line with the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.4.5.2: *"In some circumstances an analysis based on changes from baseline will be more efficient and powerful than comparison of final values"*) actually led to a great reduction of heterogeneity, for the continuous outcome, as compared to the analysis of final scores. In order to compute the change-from-baseline standard deviations we followed the method described in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 16.1.3.2), assuming a correlation coefficient of 0.5.

Unit of analysis issues

Cluster-randomised trials

In cluster-randomised trials, groups of individuals rather than individuals are randomised to different interventions (Higgins 2011). We planned to include cluster-randomised trials only when the effects of clustering were taken account of. However, we found no such cases.

Cross-over trials

Cross-over trials are trials where all participants receive both the control and intervention treatment but in a different order. The major problem is a carry-over effect from the first phase to the second phase of the study, especially if the condition of interest is unstable (Elbourne 2002). As this is the case with panic disorder, randomised cross-over studies were eligible but we only used data up to the point of the first cross-over.

Studies with multiple treatment groups

For both pair-wise and network meta-analyses, where a study involved more than two treatment arms, especially in the case of dismantling studies, we combined arms as long as they could be regarded as subtypes of the same psychological therapy under review. When arms could not be regarded as if in each of them a different subtype of the same intervention was administered, we compared each arm with the common comparator separately.

When such a situation occurred, we subdivided the common comparator arm for pairwise meta-analyses (for example, we halved the sample size and the number of responders of that arm for dichotomous outcomes; for continuous outcomes, the mean and SD will remain the same but we halved the number of patients included. The common comparator was not subdivided for NMA.

Dealing with missing data

We tried to contact the study authors for all relevant missing data.

Dichotomous outcomes

We calculated the proportion of remissions and responses using an intention-to-treat analysis (ITT) following the principle 'once randomised always analysed'. To this end, we assumed all randomised patients for which outcome data were not available to be non-responders. This assumption has been used in two previous NMAs (comparing antidepressants and antimanic drugs) and has proven to be a sensible assumption (Spineli 2013). We applied the same principle to short and long-term outcomes. When dichotomous outcomes were not reported but the means and standard deviations on a panic disorder scale were reported, we calculated the number of responding or remitted participants according to a validated imputation method (Furukawa 2005). In order to check the reliability of imputed data, we used the ANOVA intraclass correlation coefficient (ICC) to calculate agreement between reported and imputed data (absolute numbers of remitters and responders) whenever they were calculated on the same scale. The ANOVA ICC was 0.81 (0.58 to 0.93) for short-term remission and 0.99 (0.94 to 1.000) for short-term response, showing an excellent correlation between reported and imputed data.

Continuous outcomes

We performed an 'available cases analysis' in which outcomes were analysed on the basis of a pre-post change. Where change scores were not reported but baseline and endpoint data were available (including patients with either a final assessment or a LOCF to the final assessment as reported in the original report), we calculated change scores and entered them in the analyses.

Missing statistics

When only P or standard error (SE) values were reported, we calculated standard deviations (SDs) (Altman 1996). If none of these values were available, and in the absence of supplementary data after requests to the authors, we calculated the SDs according to a validated imputation method (Furukawa 2006).

Assessment of heterogeneity

Pairwise meta-analyses

For each direct comparison, we calculated the Chi^2 test and I^2 statistic in order to detect the presence of heterogeneity and, respectively, assess its degree. I^2 provides an estimate of the percentage of variability in effect estimates that is due to heterogeneity rather than chance alone (Higgins 2003). We interpreted I^2 values according to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011), section 9.5.2. We also report τ^2, the between study variance in random-effects model meta-analysis. We also used visual inspection of the forest plots in order to investigate the presence and nature of statistical heterogeneity.

Network meta-analysis (NMA)

An assumption underlying NMA is that effect modifiers are similarly distributed across comparisons in the network. That means that an effect modifier should be similar in AB and BC trials in order to obtain a valid AC estimate. Equivalent formulations of the transitivity assumption are presented in Salanti 2012. In order to verify this assumption, for each comparison we compiled a table of important trial and patient characteristics and visually inspected the similarity of factors we considered likely to modify treatment effect. We also assessed the inclusion and exclusion criteria of every trial in the network to ensure that patients, trial protocols, etc. were similar in those aspects which might modify the treatment effect.

Lack of transitivity can be manifested in the data as disagreement between direct and indirect evidence (Caldwell 2005; Lu 2004; Lumley 2002). This can be evaluated statistically by contrasting the direct and the indirect estimates and calculating a test within each closed loop (Bucher 1997; Salanti 2009). The percentage of inconsistent loops in the network is reported. We examined further the data of loops that appeared particularly inconsistent. As this approach does not provide an omnibus test and is associated with multiple testing we also employed other approaches to make inferences about the statistical inconsistency. More precisely, we performed a design-by-treatment interaction test (Higgins 2012). When a small amount of inconsistency was found, we incorporated this in the estimation by fitting inconsistency models (Higgins 2012; Lu 2004).

Assessment of reporting biases

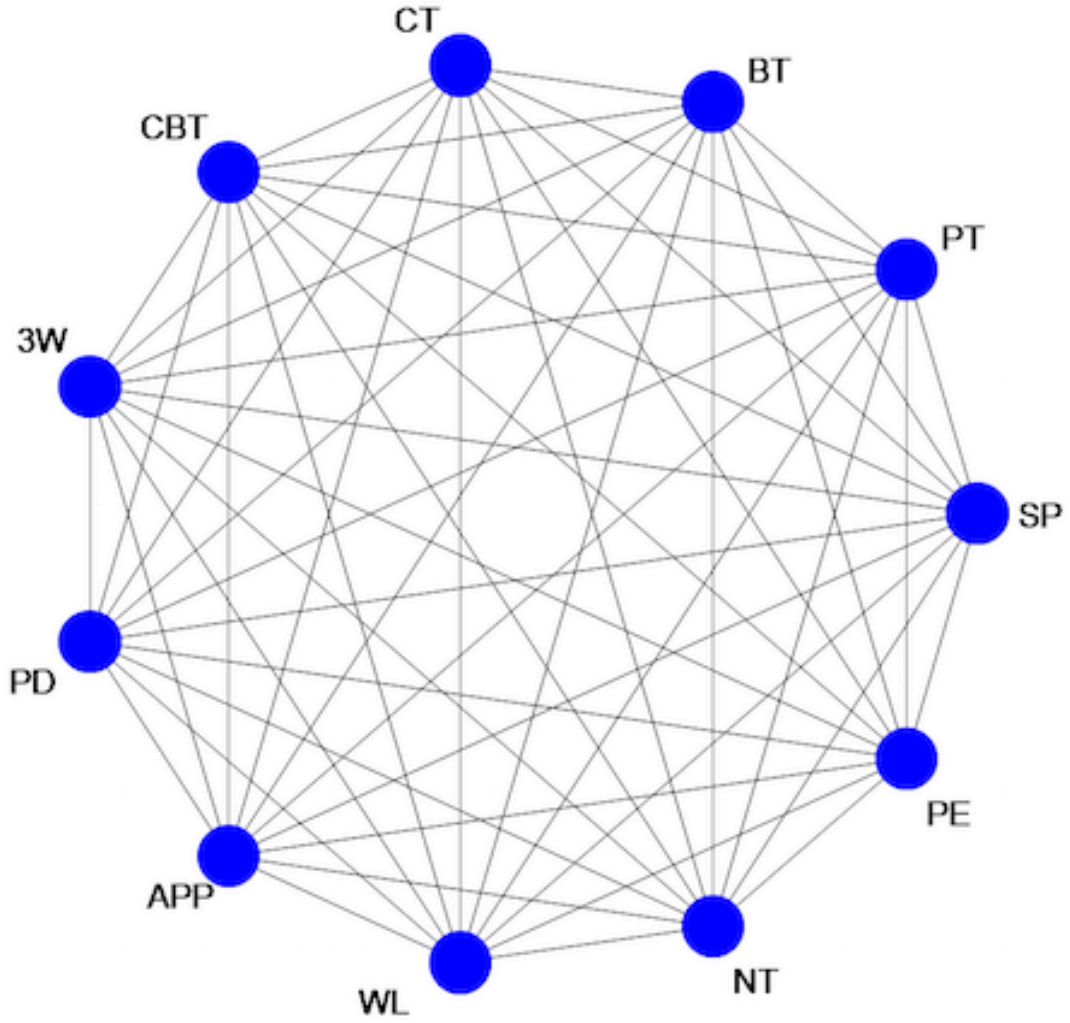
We examined the funnel plots for those pairwise comparisons for which at least 10 studies were available. We investigated the presence of small study effects for the primary outcomes only; along with visual inspection of the plots, we formally examined whether the association between estimated intervention effects and the study size was greater than it might have been expected to occur by chance.

Data synthesis

Main planned comparisons

The present study is a network meta-analysis and therefore aims to compare all the listed interventions and control conditions against one another in terms of the listed primary and secondary outcomes. In the network, each node represents an experimental or control condition; comparisons explored in included trials are represented by lines connecting the nodes. Ideally, the network should consist of 11 nodes, each connected with all the others, meaning that all the listed interventions and each possible comparison among them has been directly explored in at least one included trial (see Figure 1). Please refer to Appendix 3 for details about the software used for the analyses described below.

Figure 1. Ideal network of included treatment and control conditions.



Pairwise meta-analyses

For each available comparison explored by at least two trials, we performed a pairwise meta-analysis in order to provide overall estimates of treatment effect. Since we expected some clinical heterogeneity between studies, we planned to use a random-effects model to incorporate the assumption that the different studies are estimating different, yet related, treatment effects (Higgins 2011). We therefore calculated an 'average' treatment effect across the studies for each available comparison. For dichotomous outcomes, we calculated the average odds ratio with the 95% CI; for continuous outcomes we calculated the average SMD (or the MD if all trials use the same scale) with the 95% CI. Studies with zero events in all arms (as in the case of short-term dropouts) were not included in the analyses.

In order to have comparable results with the NMA (see below), beside performing standard pairwise meta-analyses, we also performed the analyses assuming a common heterogeneity standard deviation across all comparisons. This way, all pairwise meta-analyses were essentially analysed as random-effects (they include uncertainty due to heterogeneity), even for those comparisons only being reported by one study.

For this review, the results of pairwise comparisons are part of the more complex network meta-analyses. However, in order to better show the available 'direct' evidence, forest plots are presented for pairwise comparisons when at least 10 studies are available.

Network meta-analysis (NMA)

An indirect comparison allows an estimate of the effect of treatment B relative to treatment A via a common comparator C by statistically combining the summary effects from 'A versus C' and 'B versus C' studies (Caldwell 2005; Glenny 2005). A NMA combines direct and indirect evidence across a network of studies to make inferences regarding the relative effectiveness of multiple interventions.

A NMA is only possible for a connected set of treatments. A network diagram is constructed for our primary and secondary outcomes in order to evaluate the extent to which treatments are connected.

We conducted a random-effects model NMA, taking into account the correlations induced by multi-arm trials (Lu 2004; Salanti 2008; White 2012). For each comparison, an average effect estimate along with its 95% confidence interval (CI) is reported.

Besides yielding relative treatment effects for each comparison, a NMA allows an estimate of the relative ranking of treatments. To rank the treatments according to each outcome accounting for the uncertainty in the treatment effects, we used the surface under the cumulative ranking curve (SUCRA) (Salanti 2011). The absolute ranks of the treatments per outcome is presented using 'Rankograms' that visually show the distribution of ranking probabilities (Salanti 2011). NMA models typically employ a single heterogeneity parameter. We reported it and, for dichotomous outcomes, we judged its magnitude against the distribution of values typically found in Cochrane reviews, as presented in Turner 2012.

Subgroup analysis and investigation of heterogeneity

Subgroup and meta-regression analyses are often exploratory in nature and should be interpreted cautiously. Firstly, because these analyses often involve multiple analyses, they may yield false positive results; secondly, because these analyses lack power and are more likely to result in false negative results. Keeping in mind the above reservations, we performed meta-regression analyses to investigate, for the first of the primary outcomes only (short-term remission of panic disorder with or without agoraphobia), the following candidate explanatory variables.

- Year of publication (measured as a continuous variable) as a general proxy for various aspects (e.g. trial quality, definition of diagnosis and outcomes).
- Mean number of treatment sessions: fewer than four sessions, from four to 12 sessions, more than 12 sessions. Considerable differences exist in the number of treatment sessions between studies. It seems reasonable to expect this variability to yield some degree of heterogeneity.
- Therapist training: therapist with or without formally recognised specific training in the type of psychological therapy administered.
- Percentage of patients with agoraphobia: measured as a continuous variable.
- Percentage of patients with depression: measured as a continuous variable. We explored this variable in order to investigate if a psychological intervention specifically designed for panic disorder is less effective in patients with depressive comorbidity.
- Percentage of patients on drug treatment: measured as a continuous variable. Since we were not including studies exploring combined therapies, drug-treated patients, when included, were often those who meet the diagnosis of panic disorder despite being on psychopharmacologic treatment. By considering such patients as being 'drug-resistant', we may have expected them to have a poorer outcome; however, since there is evidence that combined therapies are more effective than psychological therapies alone in the short term (Furukawa 2007), we could also have expected that such patients had a better outcome compared with patients who were not on drug treatment.

Sensitivity analysis

The process of undertaking a systematic review and meta-analysis involves a sequence of decisions, some of which are somewhat arbitrary or unclear (Higgins 2011). A sensitivity analysis is a repeat of the primary analysis, substituting alternative decisions or range of values for decisions that were arbitrary or unclear. We planned to perform the following sensitivity analyses for the first of the primary outcomes only (short-term remission of panic disorder with or without agoraphobia).

- Restrict the inclusion in the analyses only to studies considered to be at low risk of selection and detection bias (i.e. adequate allocation sequence generation, adequate allocation concealment, blinding of assessor).
- Exclude from the analyses group therapy trials.
- Exclude from the analyses trials in which a concomitant pharmacotherapy is allowed.
- Exclude from the analyses trials in which drug therapy is not stabilised*.
- For pairwise meta-analyses, use a fixed-effect model instead of a random-effects model.

(*) Drug therapy was considered stabilised when: 1) drug administration remained stable before randomisation (for at least four weeks in the case of antidepressants and for at least two weeks in the case of benzodiazepine and other drugs), and 2) patients were asked to avoid any drug therapy change for the whole duration of the study.

'Summary of findings' tables

Aiming to summarise the results in a way that could be as 'clinically informative' as possible, we originally planned to present the main results of pairwise meta-analyses in three 'Summary of findings' tables.

In a first table we planned to present the NMA results of the comparison between the psychological therapy that would have ranked first versus the no treatment condition (NT) in order to show the effects of the supposedly most effective treatment when compared to no intervention at all.

In a second table we planned to present the NMA results of the comparison between the psychological therapy that would have ranked first versus supportive psychotherapy (SP) in order to show the effects of the supposedly most effective treatment when compared to a non-specific psychological intervention.

In a third table we planned to present the NMA results of the comparison between the psychological therapy that would have ranked first versus the one that would have ranked second, in order to show the magnitude of the effect sizes across the two active interventions representing the supposedly most viable therapeutic options.

Since the 'wait list' and 'attention or psychological placebo' conditions are useful comparators for clinical trials, but do not represent treatment options in a 'real' clinical setting, we considered the choice of using the NT and SP conditions as comparators to be more clinically informative.

As a post-hoc decision, we decided to add an extra SoF table to summarize the overall results of network meta-analyses by presenting the ranking of treatments yielded by these analyses for each outcome. We came to this decision because we found the simple presentation of pairwise comparisons, singularly taken, insufficient to adequately depict the overall complexity of this type of analyses. SoF table formats for NMAs are currently under development by the Cochrane GRADEing Group (http://methods.cochrane.org/gradeing/research), so we adapted the standard SoF for pairwise comparisons in order to present treatment hierarchy.

All the presented Sof tables include an assessment of the quality of evidence obtained by following the approach proposed by the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) Working Group. This approach consists in rating the quality of evidence according to study design (RCT or observational studies) and other five factors: risk of bias of the included studies, consistency of results, directness of evidence, precision of results and presence of publication bias. It must be noted that at the time of writing, standard GRADE tools (usually employed to assess the quality of evidence in pairwise meta-analyses) were not yet developed for NMA, where many comparisons, each with its own quality of evidence, contribute to the overall quality with different weights. Therefore, our assessments of the quality of evidence for the results of network meta-analyses were implemented by adapting the GRADE tools to this type of analysis, in line with the methodology suggested in Salanti 2014 and with interim guidance from the Cochrane Comparing Multiple Interventions Group (Cochrane Comparing Multiple Interventions Group).

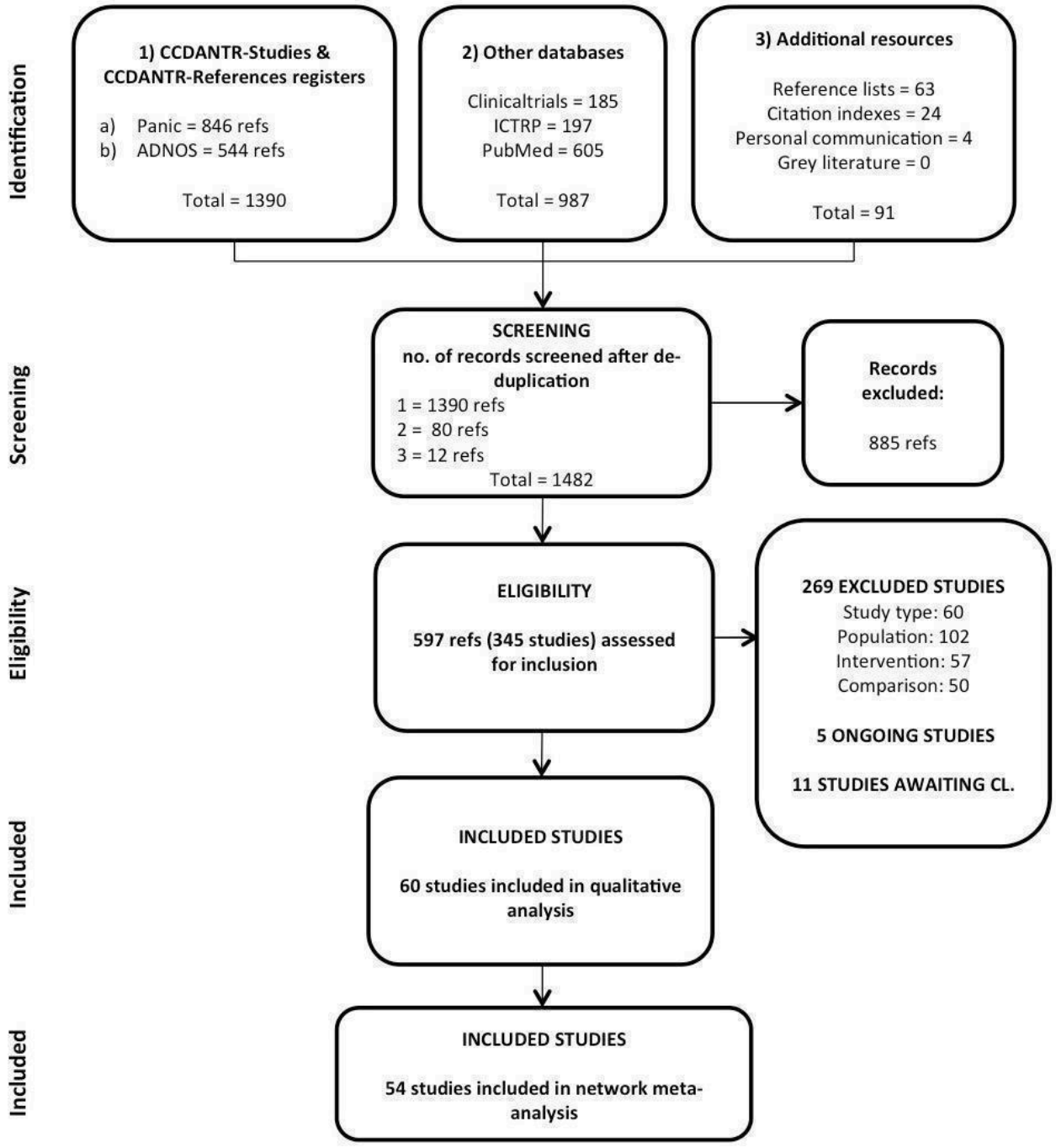
RESULTS

Description of studies

Results of the search

The number of references identified by the searches, run to March 2015, was 2468. Of these, 1482 remained after de-duplication. We excluded 885 references after assessment of the titles and abstracts. We retrieved a total of 597 full-text papers (345 studies) for full inspection. Of these 345 studies, we excluded 269 with reasons, five were ongoing trials and 11 presented too little information to be classified. We included the remaining 60 studies in the final qualitative analyses; among these, we also included 54 in the final quantitative analyses. See Figure 2 for a PRISMA flow diagram depicting the study selection process.

Figure 2. Study selection process: PRISMA flow diagram



Cohen's weighted kappa among assessors for the selection of 338 studies was 0.66 (percentage of agreement = 79.6%).

We contacted authors of 67 studies for additional information: in 25 cases we received a complete reply, in eight cases we received an incomplete reply and in the remaining 34 cases we received no reply. For six studies we have been unable to contact the author.

Included studies

We included 60 studies in this review, among which we included 54 in quantitative analyses. Five studies were published only as doctoral dissertations (Creager Berger 2001; Griegel 1995; Karekla 2004; Muncy 1991), or briefly described in a book chapter (Beck 1987; Karekla 2004).

The characteristics of the included studies can be summarised as follows (see also Characteristics of included studies).

Design

All included studies were randomised controlled trials. In only one case the study had a cross-over design (Beck 1992), although only patients firstly assigned to the control group did actually cross over in the second phase of the treatment. Only three studies had a multicentre design.

Sample sizes

In four cases the total number of randomised patients was unclear; for the remaining 56 studies, the total sample size went from 17 (Malbos 2011) to 369 patients (Gloster 2010), with a mean sample size of 60 patients (standard deviation (SD) 52). The total number of patients included in the analyses is 3021.

Setting

Apart from one single case (Hoffart 1995), which was conducted in an inpatient setting, all studies were conducted in an outpatient setting.

Participants

The presence of agoraphobia was never an exclusion criteria. Participants were therefore diagnosed with panic disorder with or without agoraphobia, and diagnosis was mostly based on DSM-III, DSM-III-R or DSM-IV; in only one case it was based on ICD-10.

Age usually ranged between 30 and 40 years. The percentage of agoraphobic patients, when specified, ranged from 18% to 100%, being above 65% in the majority of cases. In about half of the included studies participants were required to be off medication for the duration of the trial; in the remaining cases, the percentage of patients on drug treatment, when reported for the full intention-to-treat (ITT) sample, varied from 19% to 67%. In only a few studies comorbid depression was an exclusion criteria; the percentage of depressed patients in the remaining cases, when reported for the full ITT sample, varied from 7% to 52%.

Interventions

Among experimental interventions, cognitive behaviour therapy (CBT) was by far the most studied (42 of the 54 studies included in quantitative analyses), followed by behaviour therapy (13 studies) and physiological therapies (12 studies). Other psychological therapies were studied to a lesser degree: CT in six studies, SP in three studies, 3W in two studies, PD in two studies and PE in one study. In the majority of studies, the control condition was represented by a wait list (30 studies); APP was used in only three studies and NT in two.

Psychological therapy was individually administered in 18 studies, whereas a group therapy was used in 13 studies. In one case patients could receive both types of therapy even within the same study arm. In 22 studies the therapy format was not specified. The number of sessions went from 1 to 24 (the average was 10 sessions): sessions were weekly in almost every study. Each session could last from 30 to 150 minutes (average 73 minutes).

Therapists were specifically trained in the administered intervention in most of the studies (n = 37). No specific training was required in six studies. No detail about therapist training was reported in the remaining 11 studies.

Only 25 of the 54 studies included in quantitative analyses specified the percentage of patients receiving a drug therapy during the trial. Among these, 12 studies reported that patients were not receiving any drug therapy; the remaining 13 studies reported percentages from 19% to 67% (average value 45%). Drug stabilisation before study commencement was required in 27 studies, and the stabilisation period ranged from 1 to 24 weeks. Furthermore, 30 studies required patients not to change the dosages of taken medications for the entire duration of the study whereas four studies left the patients free to change dosages; in the remaining 20 studies no information on this issue were reported.

Outcomes

In terms of outcomes, we observed great variability. The most common measures were: panic frequency, Anxiety Sensitivity Index (ASI), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), Panic Disorder Severity Scale (PDSS), Mobility Inventory for Agoraphobia (MI), Fear Questionnaire (FQ), State–Trait Anxiety Inventory (STAI) and Beck Anxiety Inventory (BAI). Remission was often defined as being panic-free, although many other measures were used, such as scoring below a certain cut-off for any of the above mentioned scales, or meeting a composite index of high end-state functioning (where the set of criteria varied among different studies).

Excluded studies

We excluded a total of 269 studies because they did not meet our inclusion criteria regarding the type of study (n=60), the type of participants (n=102) or the type of interventions (experimental intervention, n=57; comparator intervention, n=50).

Among the excluded studies, 12 initially seemed to meet our inclusion criteria, but were subsequently excluded for the reasons reported in Characteristics of excluded studies.

Ongoing studies

We identified five ongoing studies. Two of these are two-arm trials exploring respectively CBT versus BT and CBT versus NT. The remaining four studies are multi-arm trials exploring, respectively: two different types of CBT versus WL; CBT versus two different types of BT; CBT versus PD versus PT; randomised CBT/PD versus chosen CBT/PD versus WL. For further details, see Characteristics of ongoing studies.

Studies awaiting classification

We identified 11 potentially eligible studies that have not yet been incorporated into the review. Details of these studies are presented in the table of Characteristics of studies awaiting classification. Apart from the case of Franklin 1990 (for which we were unable to contact the authors), all other authors have been contacted. We received a reply for two studies: Irgens 2009 (author unwilling to release full report before publication) and Richards 1997 (author himself was unable to retrieve the full paper).

Risk of bias in included studies

For details of the risk of bias judgements for each study, see Characteristics of included studies. A graphical representation of the overall risk of bias in included studies is presented in Figure 3 and Figure 4. The reporting and methodological quality of included studies was overall not good. This type of reporting has

been associated with an overestimate of the estimate of effect (Schulz 1995), and this should be considered when interpreting the results. Agreement between the two independent raters in the 'Risk of bias' assessment was overall low, ranging from 47% to 88% (weighted Kappa showed an even lower agreement, although this estimate may be negatively influenced by skewed distribution of assessments): this may be due to the relatively poor expertise of raters together with the generally low quality of reporting and to the high degree of methodological variability between studies.

Figure 3. 'Risk of bias' graph: review authors' judgements about each risk of bias item presented as percentages across all included studies.

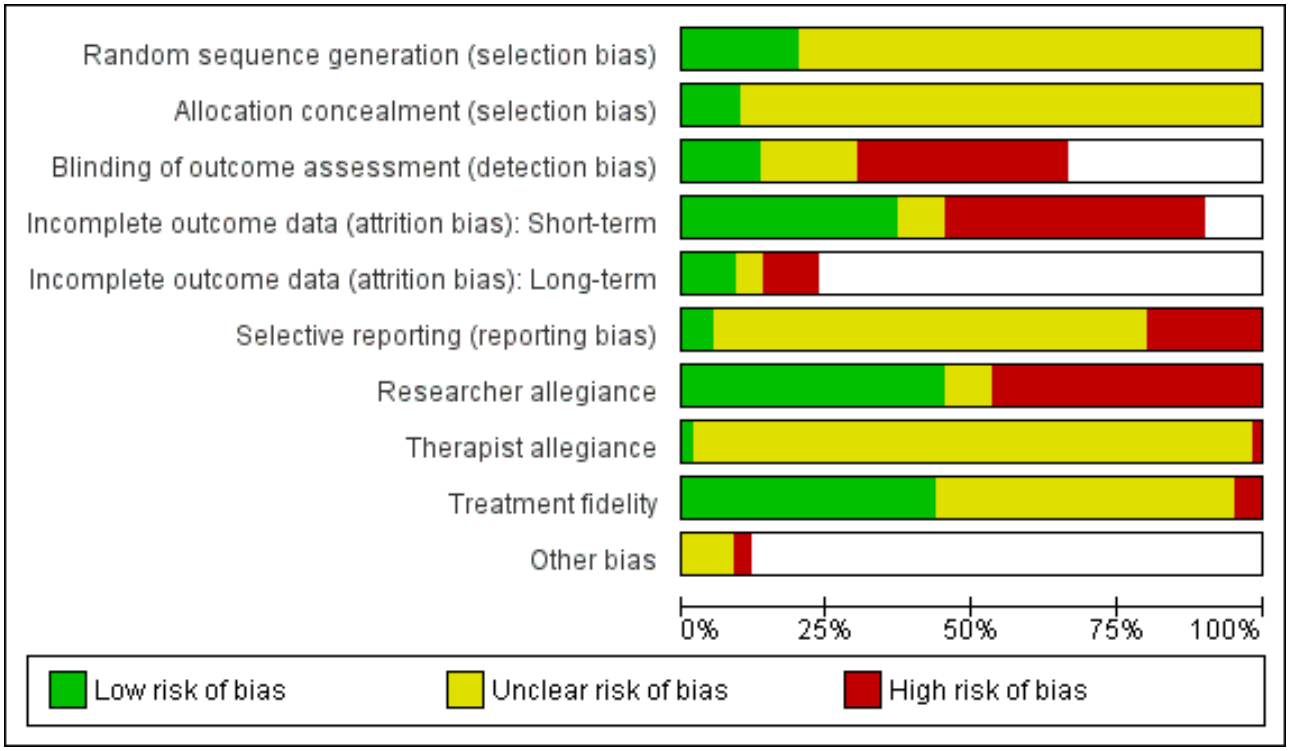


Figure 4. 'Risk of bias' summary. Note that we left the boxes empty when the RoB assessment was not applicable (e.g. in the case of incomplete outcome data in the long term when the study did not report any long term measure that could be included in the analyses).

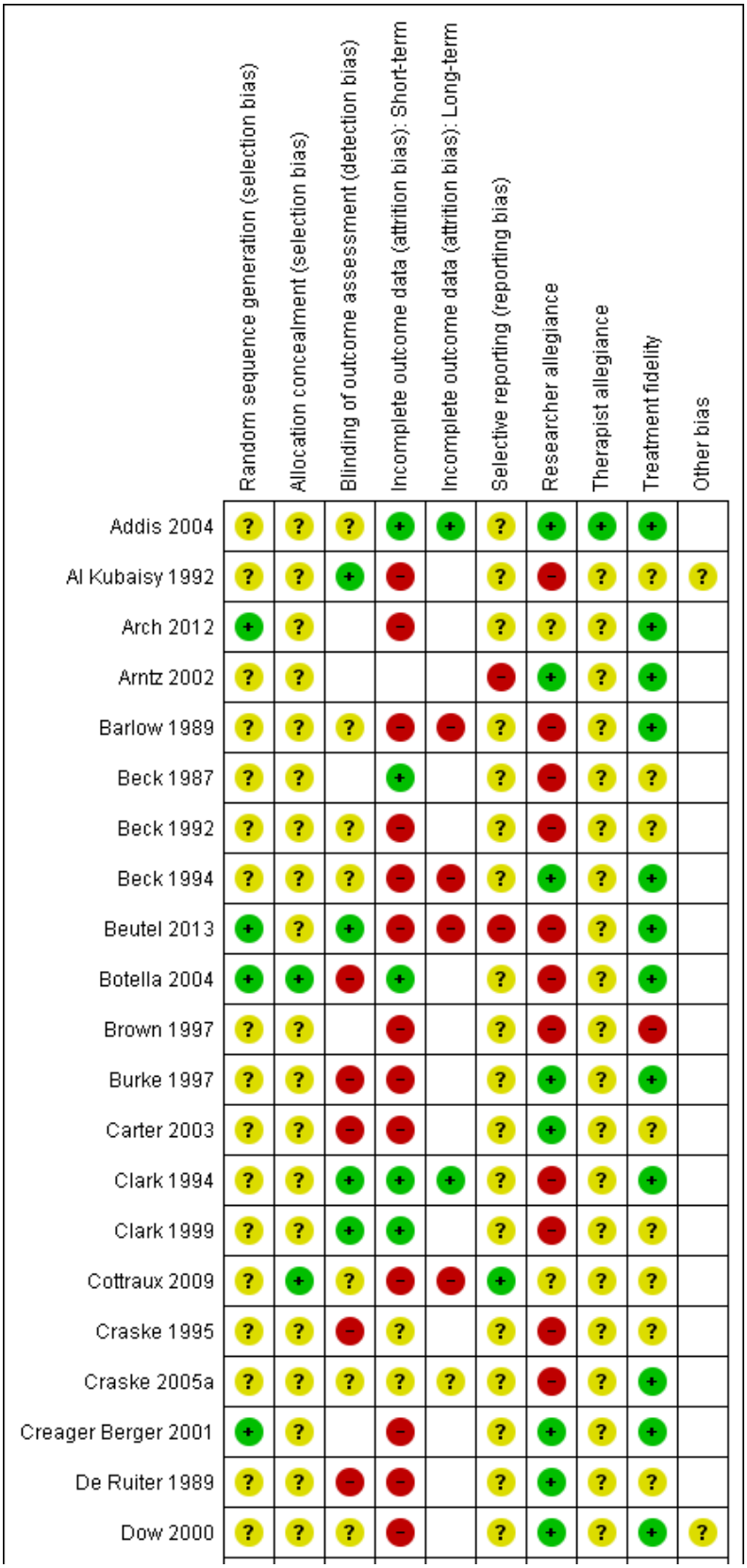


Figure 4. (Continued)

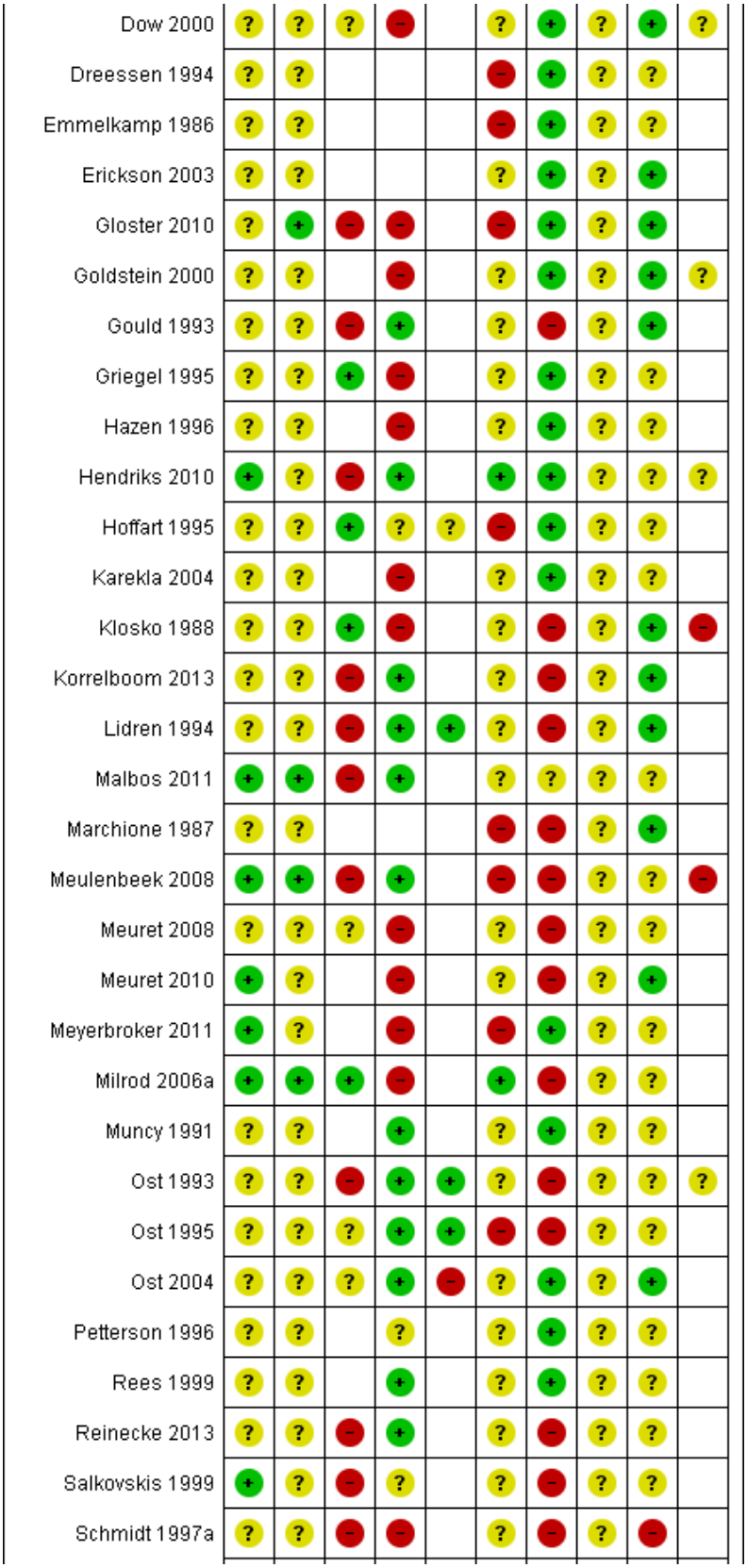
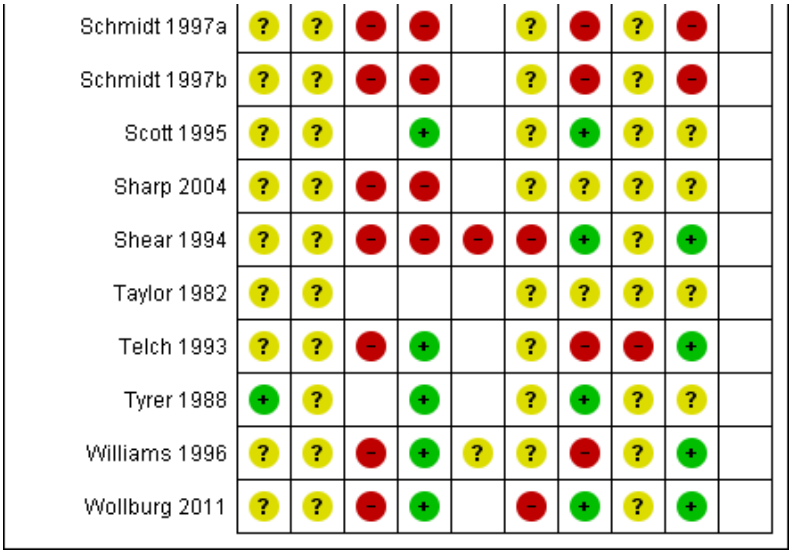


Figure 4. (Continued)



Allocation

The majority of studies did not report the methods of generating the random sequence, nor details about allocation concealment. We assessed only four studies as being at low risk of bias for both sequence generation and allocation concealment (Botella 2004; Malbos 2011; Meulenbeek 2008; Milrod 2006a). Agreement between the two independent raters in the risk of allocation bias assessment was 88% (weighted Kappa 0.47 for sequence generation and 0.29 for allocation concealment).

Blinding

We have rated the risk of bias for blinding of outcome assessment in relation to the first of our primary outcomes only (short-term remission), whenever available (either reported or imputed from continuous scale), in relation to the measure actually entered in the analyses (in this sense, our aim was not to rate study quality but rather the quality of available data). We considered eight studies to be at low risk of bias (Al Kubaisy 1992; Beutel 2013; Clark 1994; Clark 1999; Griegel 1995; Hoffart 1995; Klosko 1988; Milrod 2006a); we considered 22 studies to be at high risk of bias; 10 studies did not report enough information to make a judgement. In the remaining cases, we did not rate blinding of outcome assessment because data regarding short-term remission were not available. Agreement between the two independent raters in the risk of detection bias assessment was 47% (weighted Kappa 0.16).

Incomplete outcome data

We have rated the risk of incomplete outcome reporting when at least one relevant outcome was available (as for blinding, our aim was not to rate study quality but rather the quality of available data). Agreement between the two independent raters in the risk of attrition bias assessment was 59% (weighted Kappa 0.36) for short-term outcomes and 50% (weighted Kappa 0.12) for long-term outcomes.

Short-term

We have rated the risk of bias for incomplete outcome data assessment in relation to short-term outcomes whenever at least one of such outcomes was reported (ST-remission, ST-response, ST-dropouts, ST-improvement as measured on a continuous scale). We have rated 22 studies as being at low risk of attrition bias and 27 studies as being at high risk; five studies did not report enough information to make a judgement.

Long-term

We have rated the risk of bias for incomplete outcome data assessment in relation to long-term remission or response, whenever reported. Long-term outcome data were available in 13 studies, among which we rated five as being at low risk of attrition bias and six as being at high risk; three studies did not report enough information to make a judgement. In two cases (Cottraux 2009; Shear 1994), we did not enter long-term outcome data in the analyses because of excessive loss of data at follow-up assessments (see also Secondary outcomes).

Selective reporting

A study protocol was available for seven of the included studies (Beutel 2013; Cottraux 2009; Hendriks 2010; Meulenbeek 2008; Meyerbroker 2011; Milrod 2006a; Wollburg 2011). We rated only three studies as being at low risk of selective outcome reporting (Cottraux 2009; Hendriks 2010; Milrod 2006a). We rated 12 studies as being at high risk (assessment was sometimes possible in the absence of a study protocol, when the results of measures planned in the methods section were omitted from the study report). In all the remaining cases reporting bias could not be assessed. Agreement between the two independent raters in the risk of reporting bias assessment was 85% (weighted Kappa 0.56).

Other potential sources of bias

Researcher allegiance

In almost 50% of cases we rated studies as being at high risk of bias due to researcher allegiance, which can be considered a general proxy of various forms of bias that could affect results in favour of one or more study arms towards which authors may have a vested interest (i.e. authors may be involved in the conceptualisation of the treatment or in the developing of a treatment manual). In this sense, this source of bias can be considered analogue to sponsorship bias in studies involving pharmacological treatments. Agreement between the two independent raters in the risk of researcher allegiance bias assessment was 50% (weighted Kappa 0.34).

Therapist allegiance

We first hypothesised that therapist allegiance may constitute a risk of bias (as backed up by our own clinical sense and some literature). When we rated this item, however, the agreement was low. We went back to the original studies and found that they rarely provided enough information to make solid judgements. We therefore re-rated the risk of bias for therapist allegiance as 'unclear' when there was not enough information, which was the case for all studies except two, that is Addis 2004 and Telch 1993, rated as being respectively at low and high risk of bias. Agreement between the two independent raters in assessing the risk of this performance bias was 55% (weighted Kappa -0.02).

Treatment fidelity

We rated 26 of the included studies as being at low risk of bias with regard to treatment fidelity and three studies as being at high risk. In the remaining 31 studies, the available information was not enough to make a judgement. Agreement between the two independent raters in the risk of this performance bias assessment was 79% (weighted Kappa 0.63).

Effects of interventions

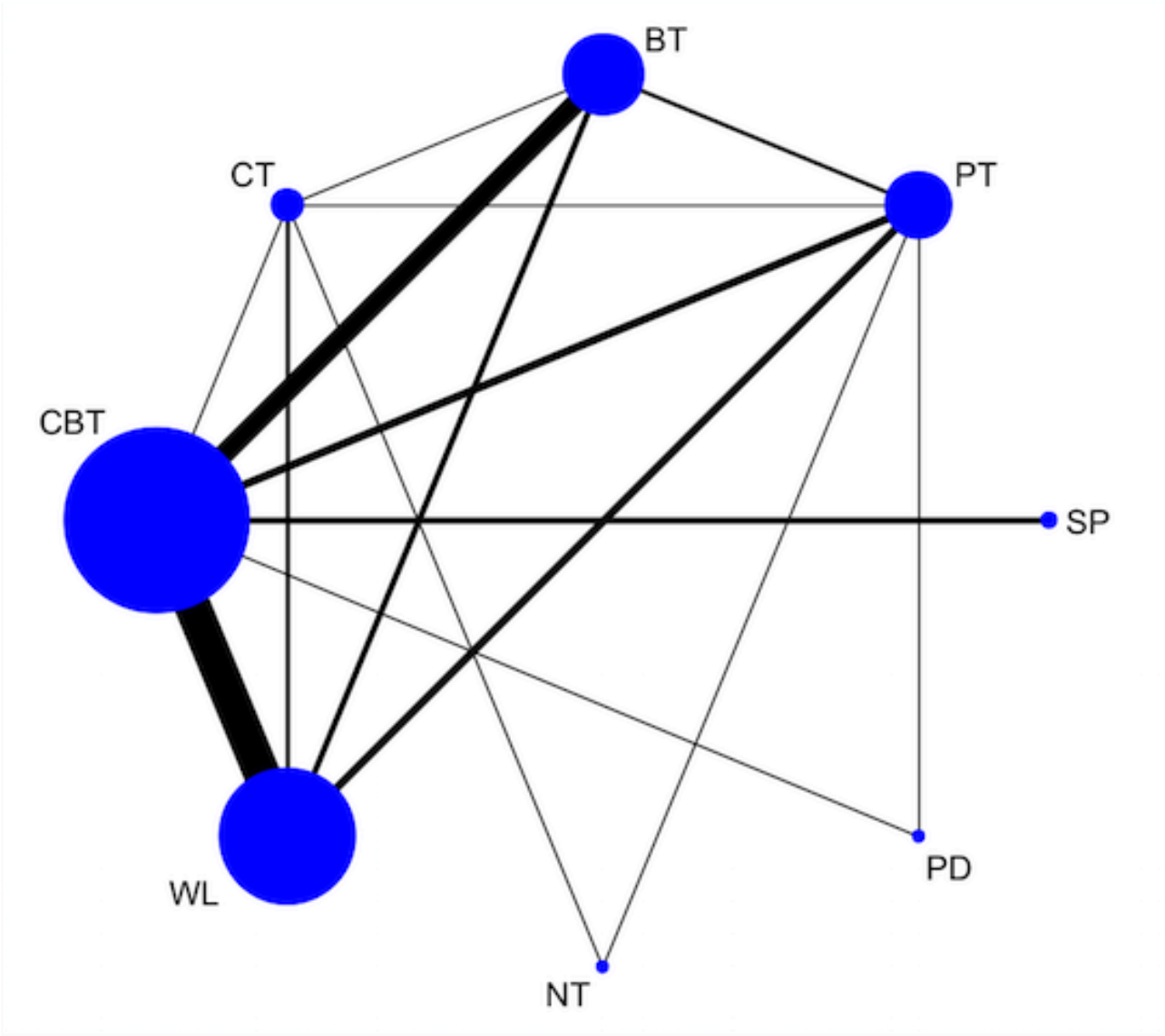
See: **Summary of findings for the main comparison** Cognitive behaviour therapy compared to no treatment for panic disorder with or without agoraphobia in adults; **Summary of findings 2** Cognitive behaviour therapy compared to supportive psychotherapy for panic disorder with or without agoraphobia in adults; **Summary of findings 3** Cognitive behaviour therapy compared to psychodynamic psychotherapy for panic disorder with or without agoraphobia in adults; **Summary of findings 4** Network meta-analysis rankings of psychological therapies for panic disorder with or without agoraphobia in adults

1. Short-term remission of panic disorder with or without agoraphobia

1.1 Network plot

Figure 5 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, short-term (ST)-remission data were available for six active and two comparison interventions. No study explored ST-remission for third-wave CBT (3W), psychoeducation (PE) and attention-psychological placebo (APP). CBT was the most studied intervention, followed by behaviour therapy (BT), physiological therapies (PT), cognitive therapy (CT) and supportive psychotherapy (SP). Waiting list (WL) was the most studied among comparator interventions. The most studied comparison was CBT versus WL, followed by CBT versus BT. The network appeared to be well connected, with the only exception being supportive psychotherapy (SP), studied only in the comparison versus CBT. Forty studies including 2491 participants contributed data to this outcome.

Figure 5. Short-term remission: network plot



1.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As explained in the methods section, in order to have comparable results with the NMA, beside standard pairwise meta-analyses, we have performed the analyses assuming a common heterogeneity standard deviation across all comparisons. The (common) heterogeneity standard deviation was estimated to be $\tau = 0.69$.

As summarised in the left part of Table 1, direct evidence was available for 15 comparisons. For seven of these comparisons there was only one study available; for the remaining eight comparisons we performed a random-effects meta-analysis. As shown in the table, only two comparisons were informed by 10 or more studies, that is CBT versus WL (18 studies) and CBT versus BT (10 studies): their forest plots are respectively presented in Figure 6 and Figure 7. Among psychological therapies, four were shown to be significantly better than WL in terms of short-term remission: PT (four studies; odds ratio (OR) 4.8, 95% confidence interval (CI) 1.4 to 17), BT (three studies; OR 8.3, 95% CI 2.3 to 25), CT (two studies; OR 8.3, 95% CI 1.6 to 50) and CBT (18 studies; OR 7.7, 95% CI 4.5 to 14.3). The comparison CBT versus BT is the only comparison among two active treatments that showed a statistically significant difference in terms of short-term remission, which was in favour of CBT (10 studies; OR 2.09, 95% CI 1.10 to 3.97).

Figure 6. Short-term remission: forest plot for the comparison WL vs CBT

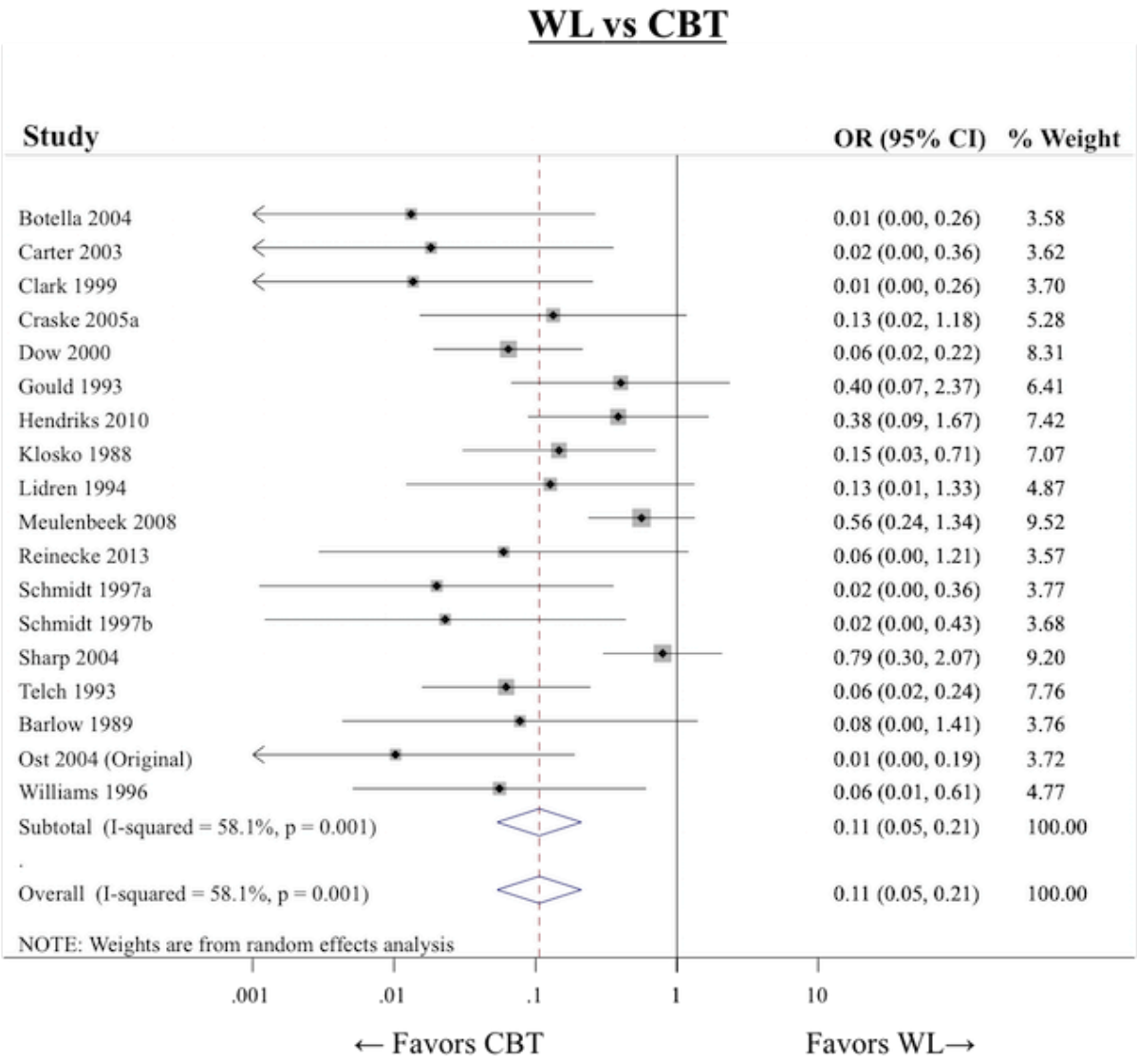
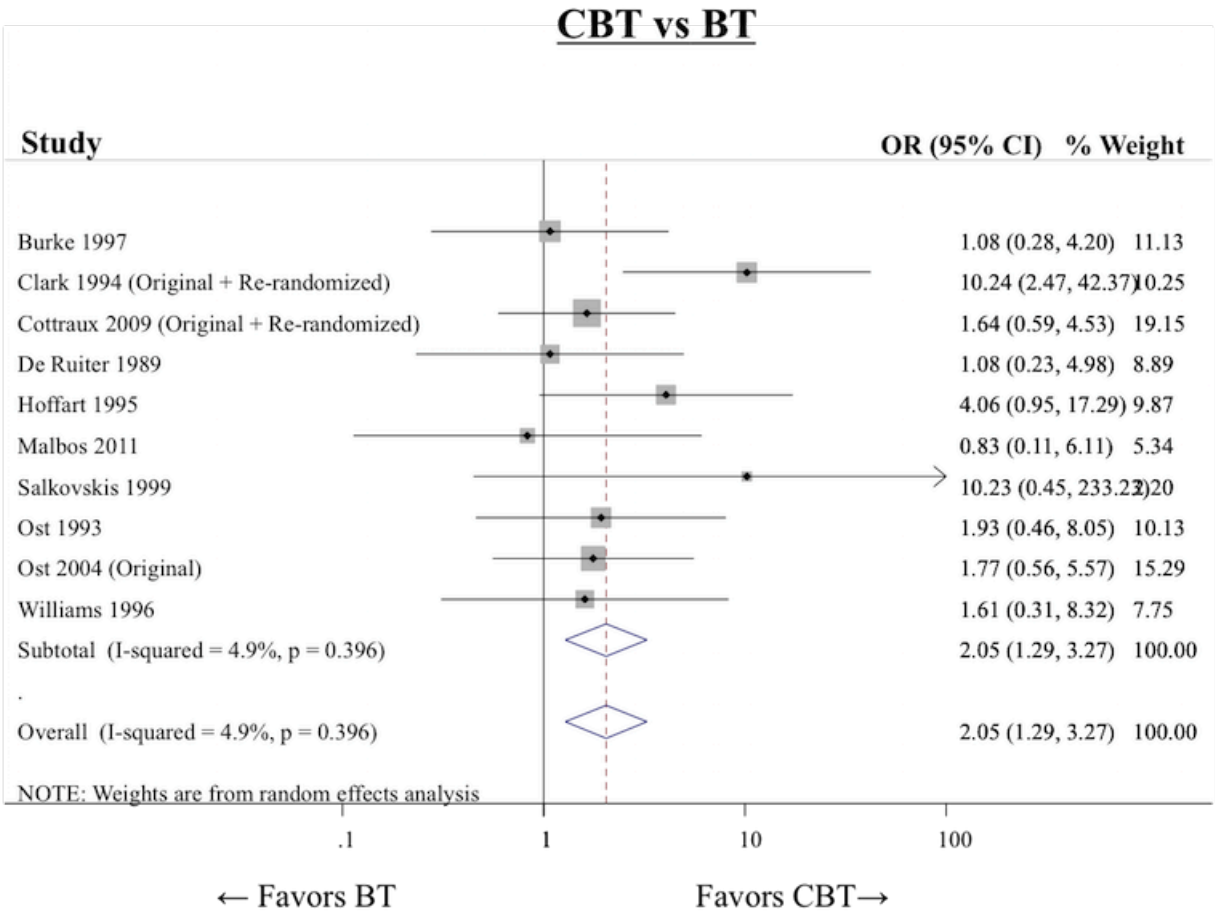


Figure 7. Short-term remission: forest plot for the comparison CBT vs BT



Heterogeneity

The I^2 values and their 95% CIs, for the comparisons reported in three studies or more, are presented in Table 2. As shown in the table, we observed the highest I^2 values in the comparisons CBT versus WL (I^2 = 58,1%) and PT versus WL (I^2 = 56%). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) these values suggest that, in these two comparisons, a moderate percentage of the observed variability in the effect estimates was due to heterogeneity rather than sampling error (chance). In the comparison CBT versus WL, heterogeneity appeared to be related to small study effect (see below); in the case of PT versus WL (four studies), heterogeneity was due to the study Griegel 1995, a three-arm trial (PT versus PT versus WL) in which no remission (i.e. panic-free status) was observed in the two active treatment arms whereas one case of remission was observed in the wait list. In this study, therefore, the unexpected OR was due to the low number of events across all arms.

Small study effects

Following the protocol, we produced funnel plots for all comparisons appearing in more than 10 studies. There were two comparisons appearing in 10 studies or more, that is WL versus CBT (Figure 8) and CBT versus BT (Figure 9). From the first funnel plot there was evidence of asymmetry. More specifically, small studies were missing in the lower right part of the funnel plot. This means that small studies comparing WL to CBT that (relatively) favour WL seemed to be missing: in other words, small studies showed CBT to be more efficacious. We performed a meta-regression for the WL versus CBT comparison, which formally confirmed the presence of a statistically significant correlation between the effect size (log odds ratio) and the variance. The contour-enhanced funnel plot for the comparison WL versus CBT (Figure 10) showed that studies were missing in the area of non-significance, thus suggesting the role of publication bias behind the small study effect (SSE). We found no evidence of asymmetry in the funnel plot for the comparison CBT versus BT.

Figure 8. Short-term remission: funnel plot for the comparison CBT vs WL

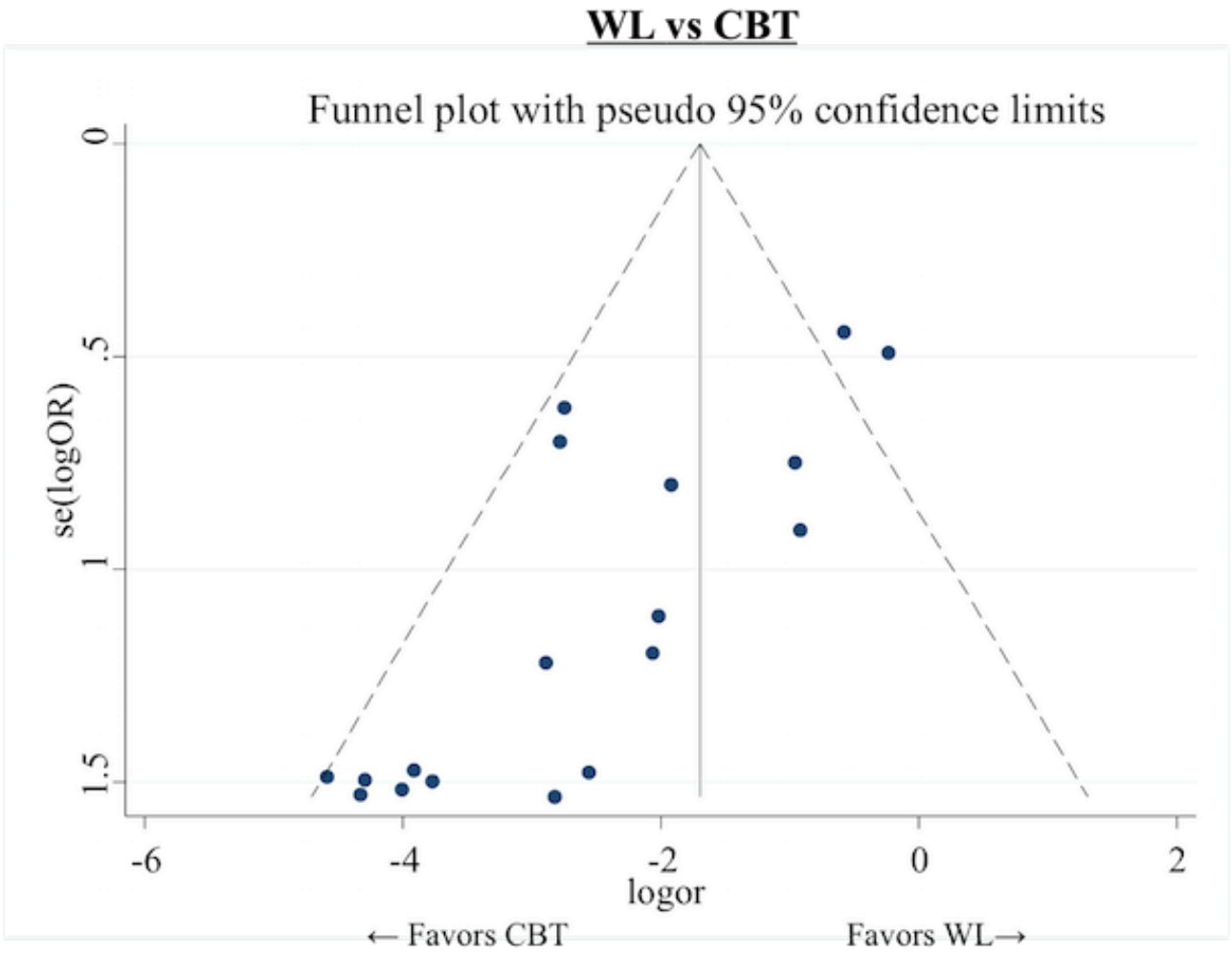


Figure 9. Short-term remission: funnel plot for the comparison CBT vs BT

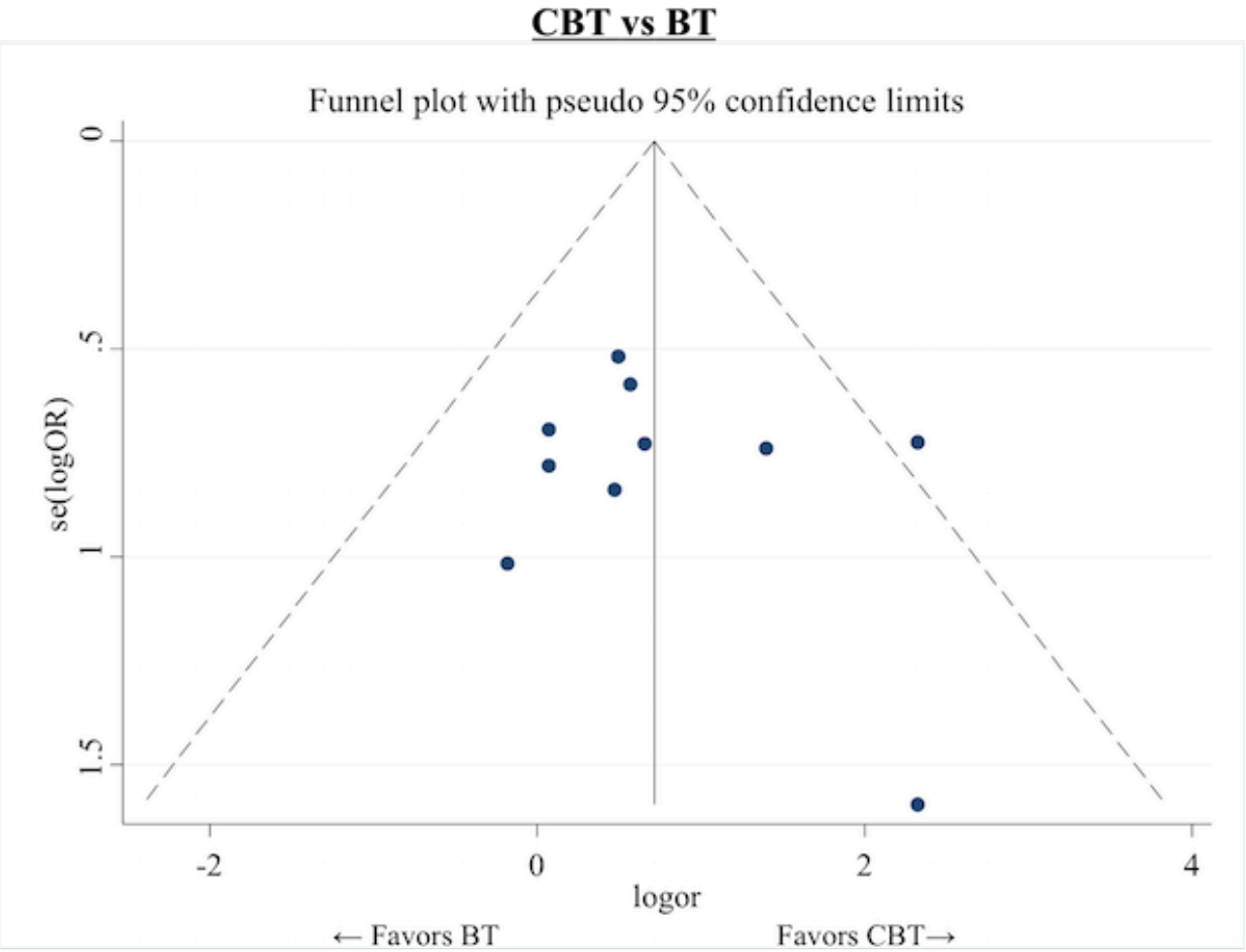
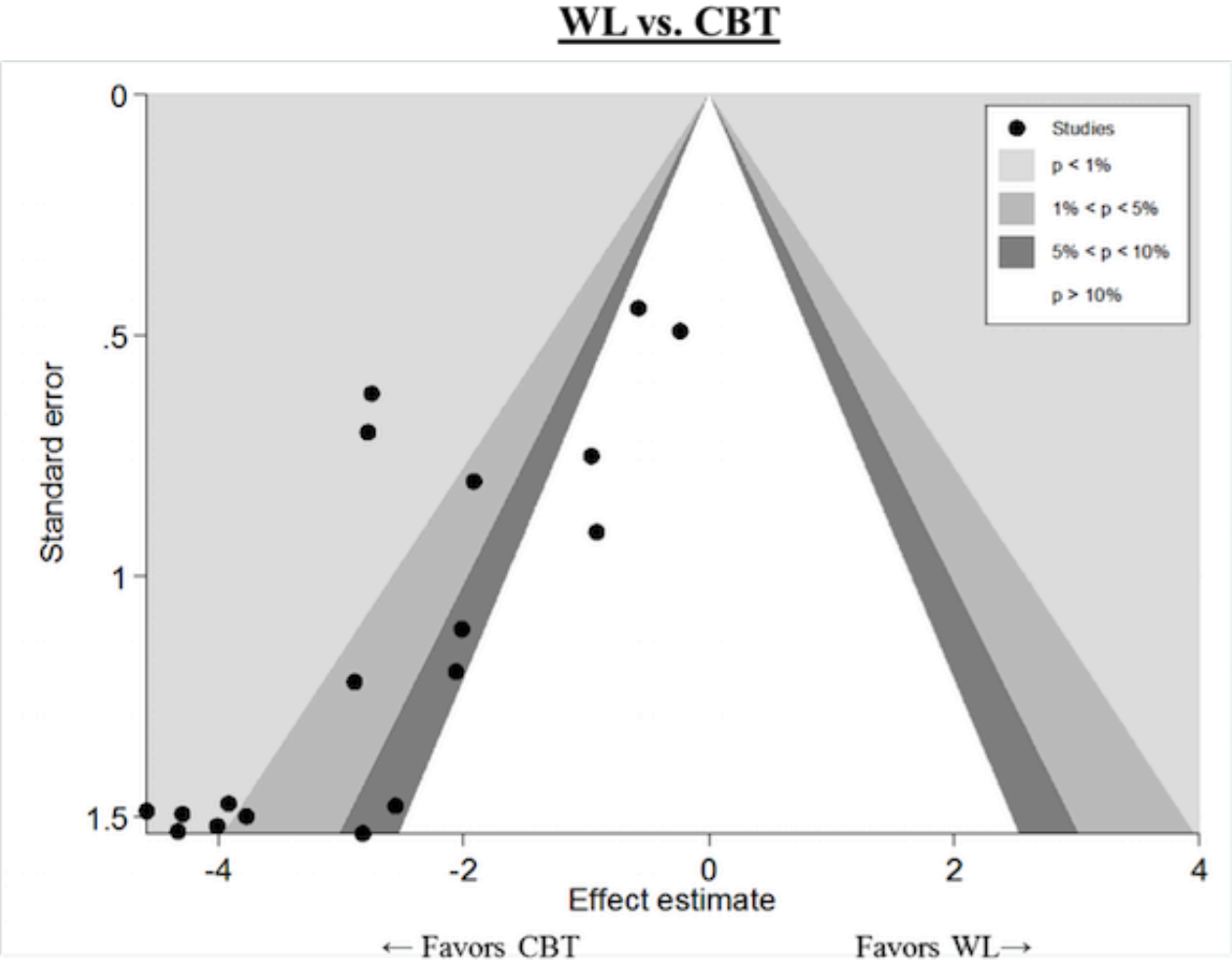


Figure 10. Short-term remission: contour-enhanced funnel plot for the comparison CBT vs WL



1.3 Network meta-analysis and its inconsistency

Network meta-analysis

As explained in the previous paragraph, it was evident from the funnel plots that there were small study effects (SSE) present in the network for the comparison CBT versus WL. We found it reasonable to assume that there were SSE in all other comparisons versus WL, even though we might not have had enough studies to see this effect. The presence of SSE implies that a simple NMA may produce biased results. For this reason we performed a network meta-analysis adjusting for SSE in studies comparing all other treatments to WL, by regressing on the variance of the study (see Discussion). We performed the network meta-analysis adjusted for SSE in WinBUGS. Thus the results are expressed in terms of credible intervals (CrI) and we use the median (instead of the mean) because the posterior distribution of the estimated odds ratios is asymmetrical.

Results of the network meta-analysis (NMA) for short-term remission, unadjusted and adjusted for SSE, are presented in Table 1. Indirect evidence could be calculated for 13 comparisons for which direct evidence was unavailable. The comparison between CBT and WL remained statistically significant also within the context of NMA, showing an OR of 3.0 in favour of CBT (95% CrI 1.5 to 6.3) in the analyses adjusted for SSE (note that the point estimate in standard NMA was higher, with an OR of 8.3). Although unadjusted NMA basically confirmed the results of pairwise meta-analyses for PT, CT and BT in the comparison versus WL, results ceased to be statistically significant in the NMA adjusted for SSE. We found supportive psychotherapy (SP) to be significantly better than WL (OR 4.5, CrI 1.3 to 16.7); however, this finding must be interpreted with caution since SP is included in the network as a node with a single connection (see Discussion). Finally, two comparisons among active treatments, that is CBT versus BT and CBT versus PT, showed a statistically significant difference in terms of short-term remission, in both cases in favour of CBT, with an OR respectively of 1.77 (CrI 1.02 to 3.11) and 1.95 (CrI 1.02 to 3.97).

Network heterogeneity and inconsistency

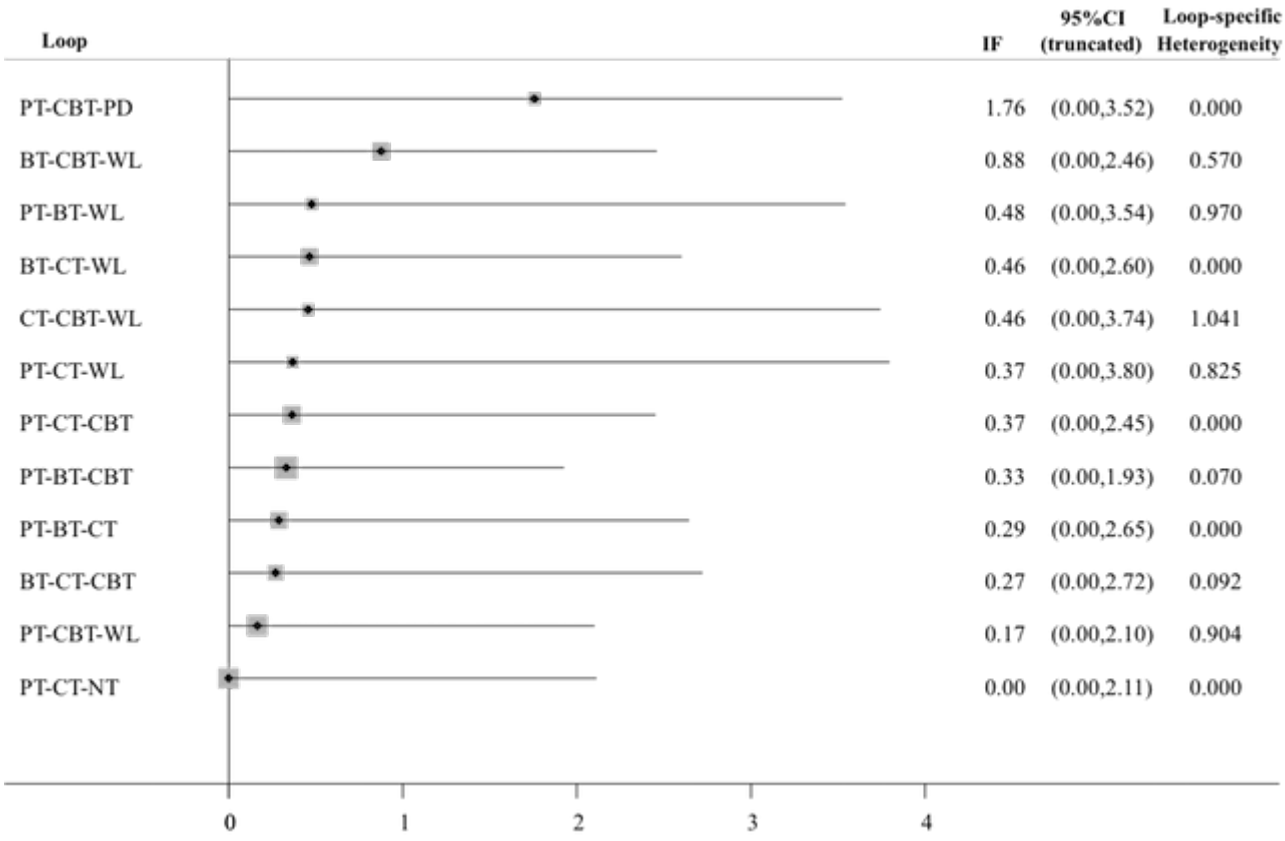
For both adjusted and unadjusted NMAs, the estimated values of heterogeneity lay well within the range of values usually found in Cochrane reviews, as presented by Turner 2012.

We compiled a table of important trial and patient characteristics including therapy duration and percentage of agoraphobic, depressed and drug treated patients. Its visual inspection showed

that those effect modifiers were similarly distributed across comparisons in the network: we therefore concluded that there wasn't important evidence against the transitivity assumption.

We compared the inconsistency factors using the loop-specific approach (where we allow the same τ for all comparisons in a loop) before and after the adjustment for small study effects. We observed no important differences and all inconsistency factors were statistically non-significant in both cases. One, however, should note that this does not constitute a proof for consistency in the network: some of the loops include few studies and the corresponding factors are estimated with much uncertainty. In Figure 11 we give all inconsistency factors for the network. As shown in the figure, we observed the highest inconsistency factor in the loop PT-CBT-PD.

Figure 11. Short-term remission: inconsistency factors for the network



The design-by-treatment interaction model provided no proof of global inconsistency in the network (Chi^2 = 4.32 with 13 degrees of freedom; P value for the null hypothesis of consistency in the network 0.98). Also, using the design-by-treatment inconsistency model we got an estimate of τ = 0.85, a value higher than the one we obtained from the consistency model. Thus, we conclude that there was no proof of inconsistency in the network. Again, this does not constitute a proof of the absence of consistency because the network was underpowered to detect any important disagreement between direct and indirect evidence.

1.4 Ranking of treatments

The ranking of treatments with respect to short-term remission, according to the surface under the cumulative ranking curve (SUCRA) value derived from NMA adjusted for small study effects, is presented in Table 3. We observed the highest rankings respectively for supportive psychotherapy, cognitive behaviour therapy and psychodynamic therapy. However, results regarding supportive psychotherapy must be interpreted with caution because, as specified earlier, SP is included in the network as a node with a single connection to the network, being compared only with CBT (three studies, OR 1.5 in favour of SP, 95% CrI 0.6 to 4).

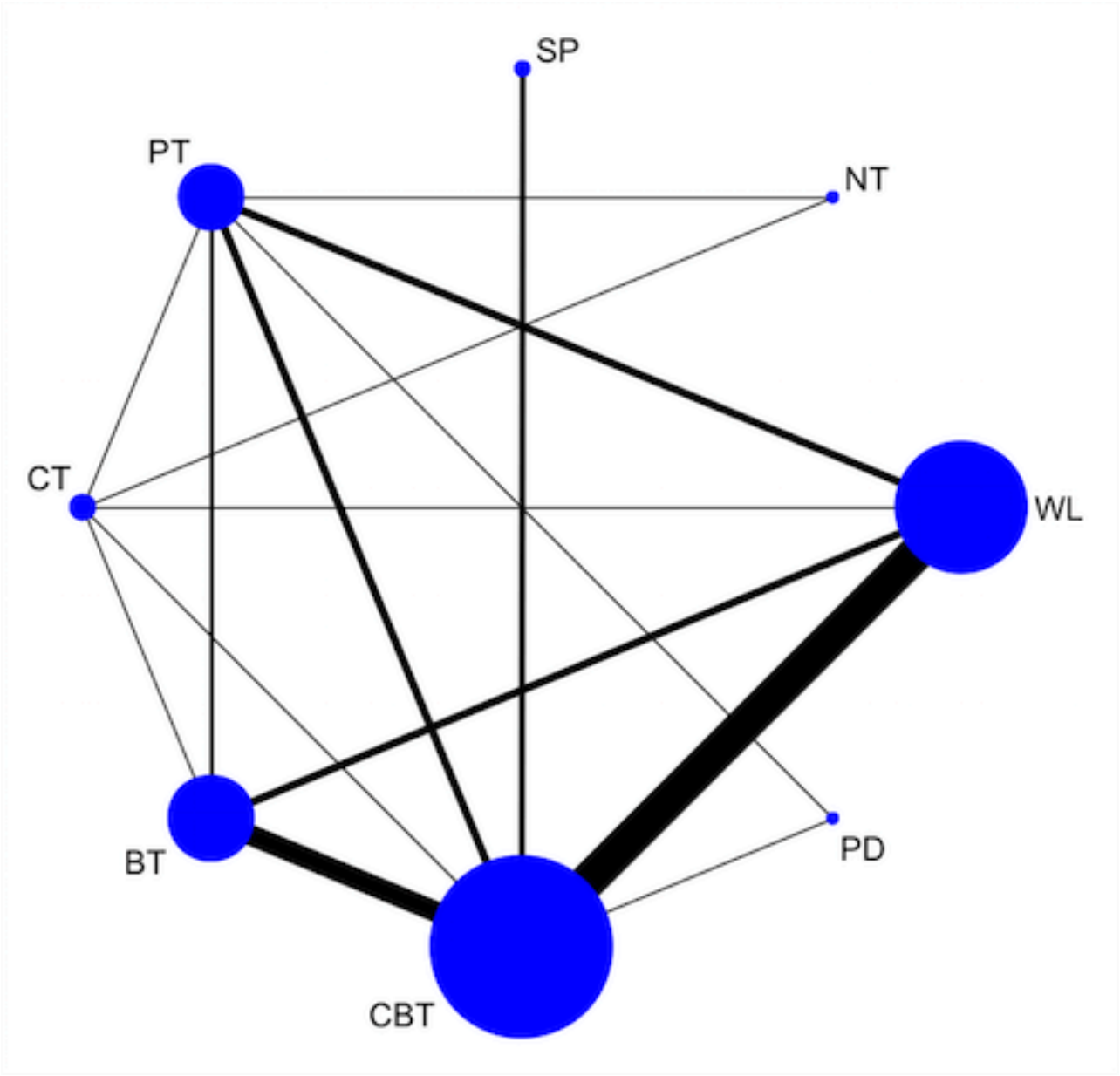
2. Short-term response of panic disorder with or without agoraphobia

2.1 Network plot

Figure 12 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. The network is similar to the one for ST-remission: as shown in the figure, ST-response data were available for six psychological therapies and two control interventions. No study explored ST-response for third-wave CBT (3W), psychoeducation (PE) and attention-psychological placebo (APP). CBT appeared to be the most studied intervention, followed by behaviour therapy (BT), physiological therapies (PT), cognitive therapy (CT) and supportive psychotherapy (SP). Wait list (WL) was the most studied among comparator interventions. The most studied comparison was CBT versus WL, followed by CBT versus BT. The network appeared to be well connected, with the only exception being supportive psychotherapy (SP), studied only

in the comparison versus CBT. Thirty-seven studies including 2240 participants contributed data to this outcome.

Figure 12. Short-term response: network plot



2.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As for short-term remission, in order to have comparable results with the NMA we have performed the pairwise meta-analyses assuming a common heterogeneity variance across all comparisons. The (common) heterogeneity standard deviation was estimated to be $\tau = 0.55$.

As summarised in the left part of Table 4, direct evidence was available for 15 comparisons. For eight of these comparisons there was only one study available; for the remaining seven comparisons we performed a random-effects meta-analysis. As shown in the table, only two comparisons were informed by 10 or more studies, that is CBT versus WL (17 studies) and CBT versus BT (10 studies): their forest plots are respectively presented in Figure 13 and Figure 14. Among psychological therapies, three were shown to be significantly better than WL in terms of short-term response: PT (four studies; OR 6.67, 95% CI 2.27 to 20), BT (four studies; OR 3.13, 95% CI 1.37 to 7.14) and CBT (17 studies; OR 5.26, 95% CI 3.23 to 20). The comparison CBT versus BT is the only comparison among two active treatments that showed a statistically significant difference in terms of short-term remission, which was in favour of CBT (10 studies; OR 1.78, 95% CI 1.0 to 3.18).

Figure 13. Short-term response: forest plot for the comparison WL vs CBT

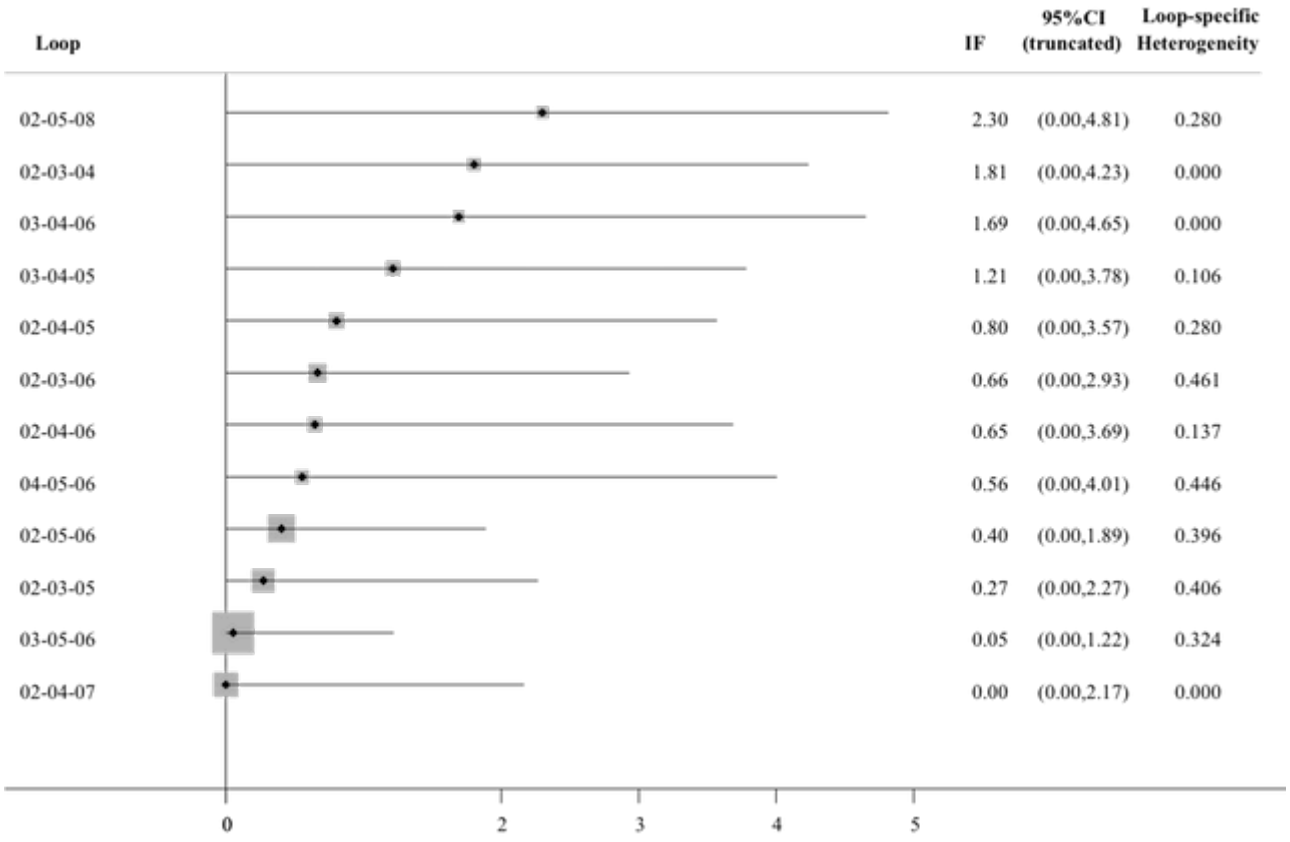
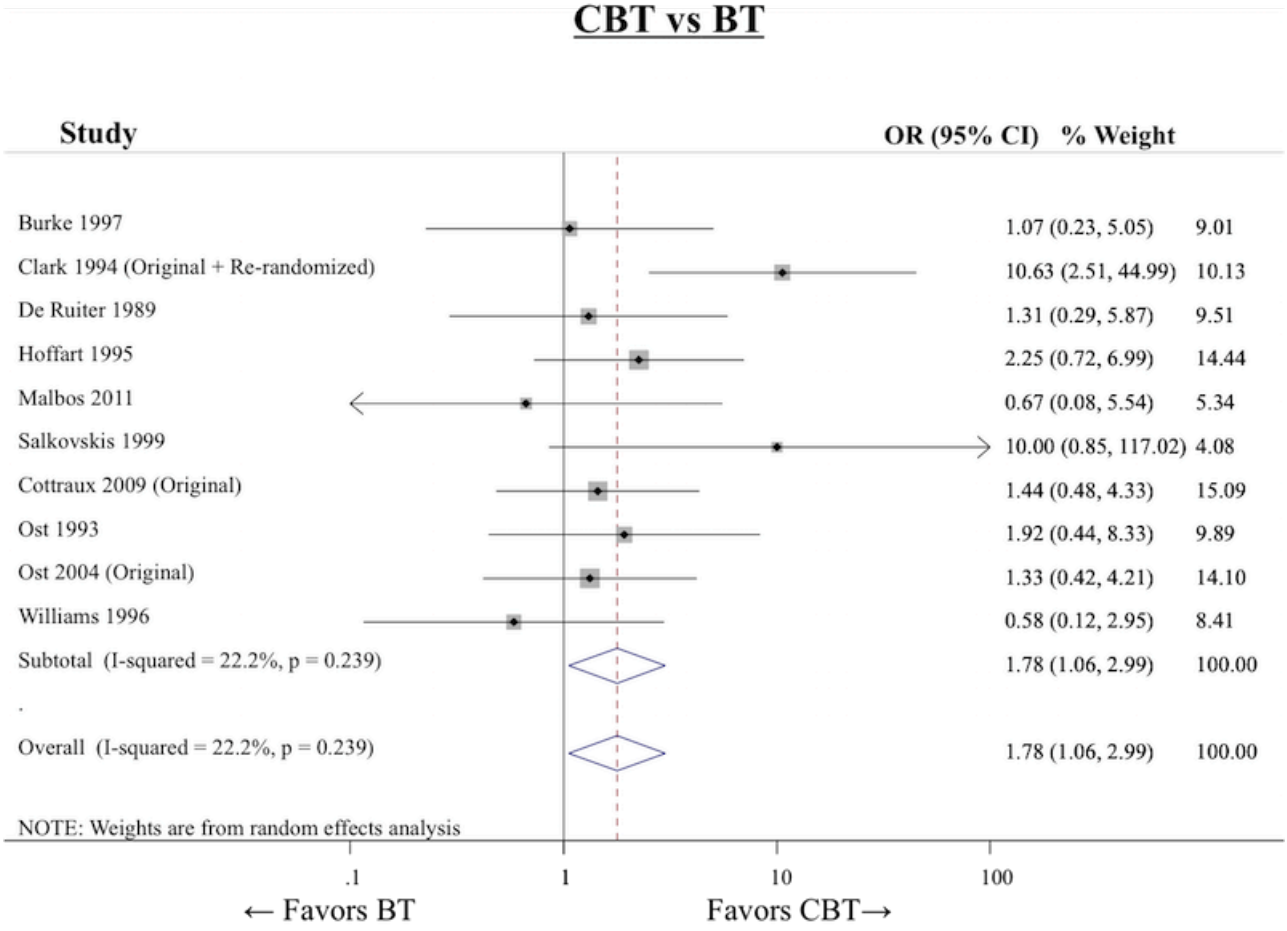


Figure 14. Short-term response: forest plot for the comparison CBT vs BT



Heterogeneity

I^2 values and their 95% CIs, for the comparisons reported in three studies or more, are presented in Table 5. As shown in the table, we observed the highest I^2 values in the comparisons CBT versus PT (I^2 = 45%) and WL versus CBT (I^2 = 39%). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) these values suggest that, in these two comparisons, a moderate percentage of the observed variability in the effect estimates was due to heterogeneity rather than sampling error (chance). In the case of CBT versus PT (four studies) we could not find a clear explanation for the observed heterogeneity; in the comparison CBT versus WL heterogeneity appeared to be related to small study effects (see below), as was the case for short-term remission.

Small study effects

Following the protocol, we produced funnel plots for all comparisons appearing in more than 10 studies. There were two comparisons appearing in 10 studies or more, that is WL versus CBT (Figure 15) and CBT versus BT (Figure 16). From the first funnel plot there was evidence of asymmetry. More specifically, small studies were missing in the lower right part of the funnel plot. This means that small studies comparing WL to CBT that (relatively) favour WL seemed to be missing: in other words, small studies showed CBT to be more efficacious. Similar to ST-remission, the contour-enhanced funnel plot for the comparison WL versus CBT (not presented) showed that studies were missing mainly in the area of non-significance, thus suggesting the role of publication bias behind the SSE. We found no evidence of asymmetry in the funnel plot for the comparison CBT versus BT.

Figure 15. Short-term response: funnel plot for the comparison WL vs CBT

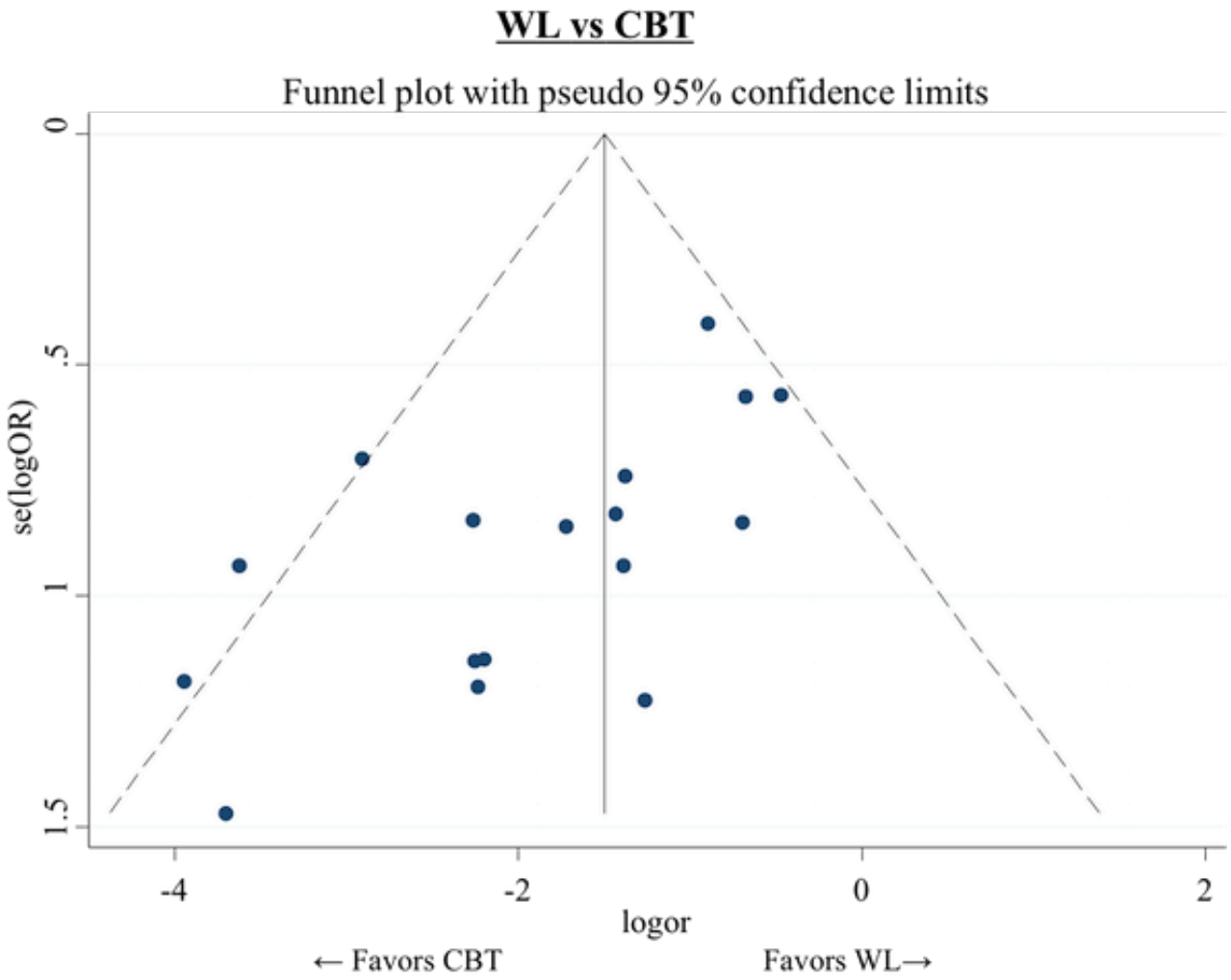
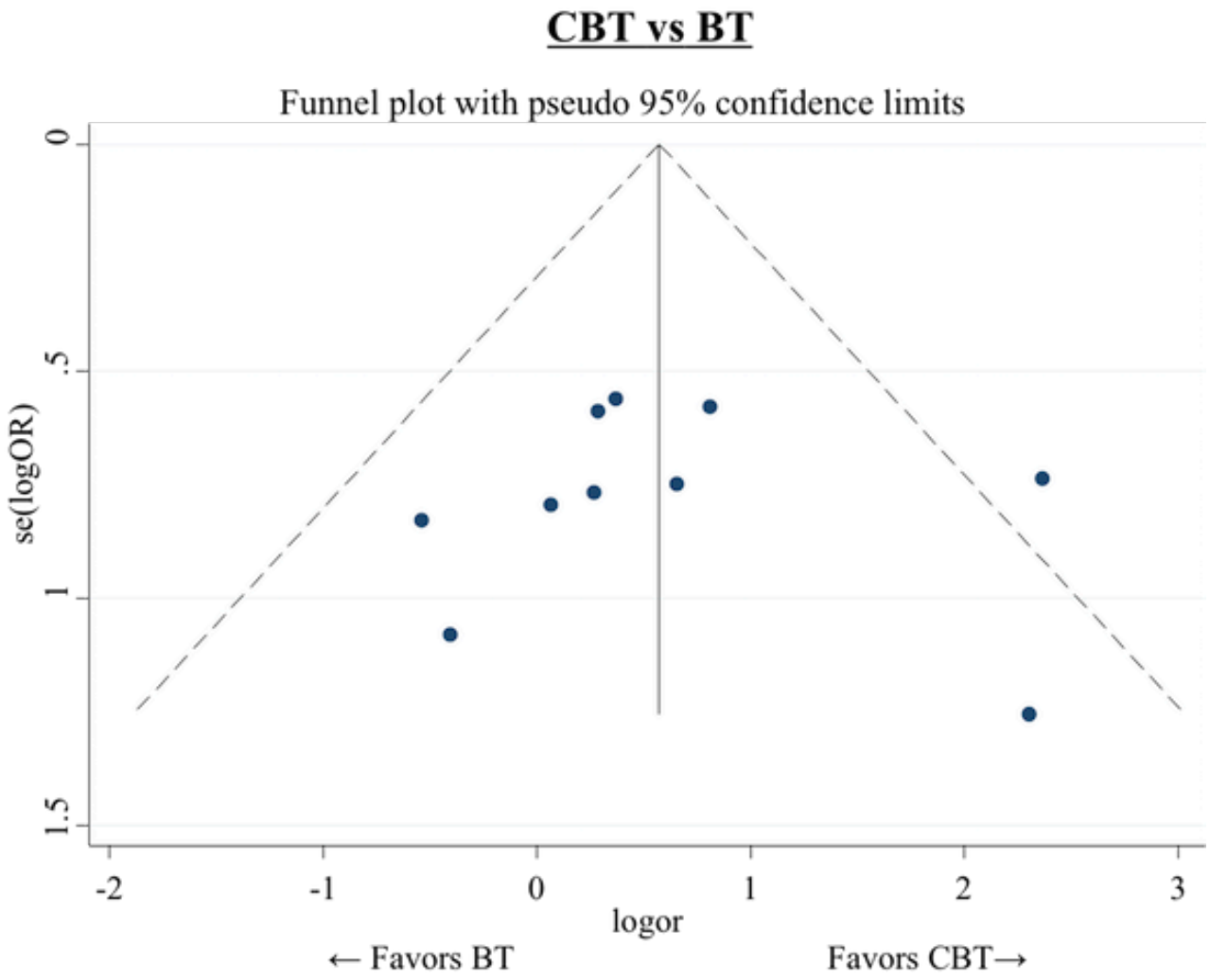


Figure 16. Short-term response: funnel plot for the comparison CBT vs BT



2.3 Network meta-analysis and its inconsistency

Network meta-analysis

As was the case for short-term remission, it was evident from the funnel plots that there were small study effects (SSE) present in the network for the comparison CBT versus WL. We found it reasonable to assume that there were SSE in all other comparisons versus WL, even though we might not have had enough studies to see this effect. For this reason we performed a network meta-analysis adjusting for SSE in studies comparing all other treatments to WL, by regressing on the variance of the study. We performed the network meta-analysis adjusted for SSE in WinBUGS. Thus the results are expressed in terms of credible intervals and we use the median (instead of the mean) because the posterior distribution of the estimated odds ratios is asymmetrical.

Results of the network meta-analysis (NMA) for short-term response, unadjusted and adjusted for SSE, are presented in Table 4. Indirect evidence could be calculated for 13 comparisons for which direct evidence was unavailable. The three comparisons CBT versus WL, BT versus WL and PT versus WL lost statistical significance within the context of NMA adjusted for SSE. The same happened for the comparison CBT versus BT. Interestingly, all cited comparisons showed a statistically significant difference in the standard NMA (central part of Table 4), but lost significance when adjusting the analyses for SSE. The only comparison that showed a statistically significant difference in the NMA adjusted for SSE was CBT versus NT (indirect evidence only), with an OR of 7.14 (95% CrI 1.25 to 50).

Network heterogeneity and inconsistency

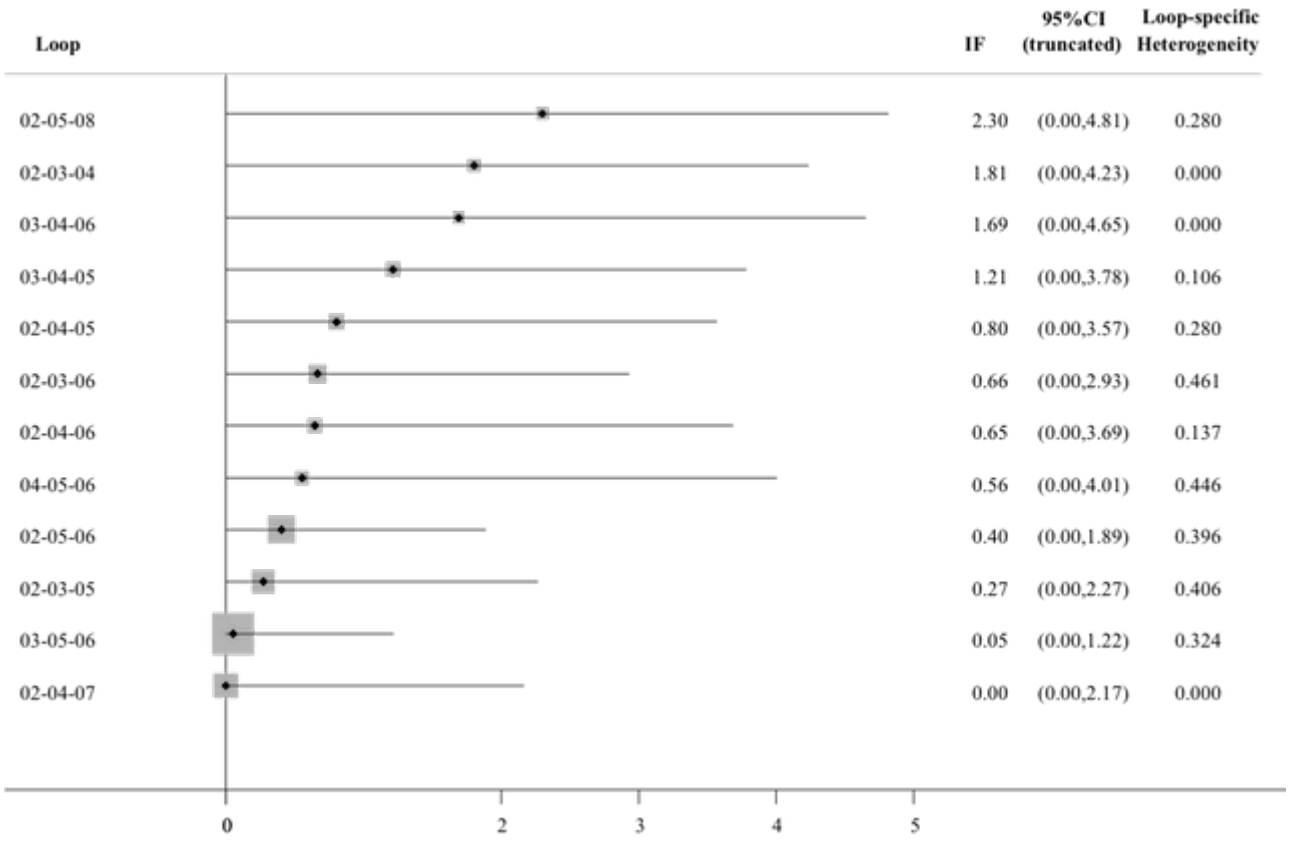
For both adjusted and unadjusted NMAs, the estimated values of heterogeneity lay well within the range of values usually found in Cochrane reviews, as presented by Turner 2012.

We compiled a table of important trial and patient characteristics including therapy duration and percentage of agoraphobic, depressed and drug treated patients. Its visual inspection showed that those effect modifiers were similarly distributed across comparisons in the network: we therefore concluded that there wasn't important evidence against the transitivity assumption.

We compared the inconsistency factors using the loop-specific approach (where we allow the same τ for all comparisons in a loop) before and after the adjustment for small study effects. We observed no important differences and all inconsistency

factors were statistically non-significant in both cases. However, as explained in section 1.3, this does not constitute a proof of consistency in the network: some of the loops include few studies and the corresponding factors are estimated with much uncertainty. In Figure 17 we give all inconsistency factors for the network. Again, as in the network for short-term remission, we observed the highest inconsistency factor in the loop PT-CBT-PD.

Figure 17. Short-term response: inconsistency factors for the network



We found no proof of global inconsistency using the design-by-treatment inconsistency model (Chi^2 = 7.74 with 12 degrees of freedom; P value for the null hypothesis of consistency in the network 0.80).

2.4 Ranking of treatments

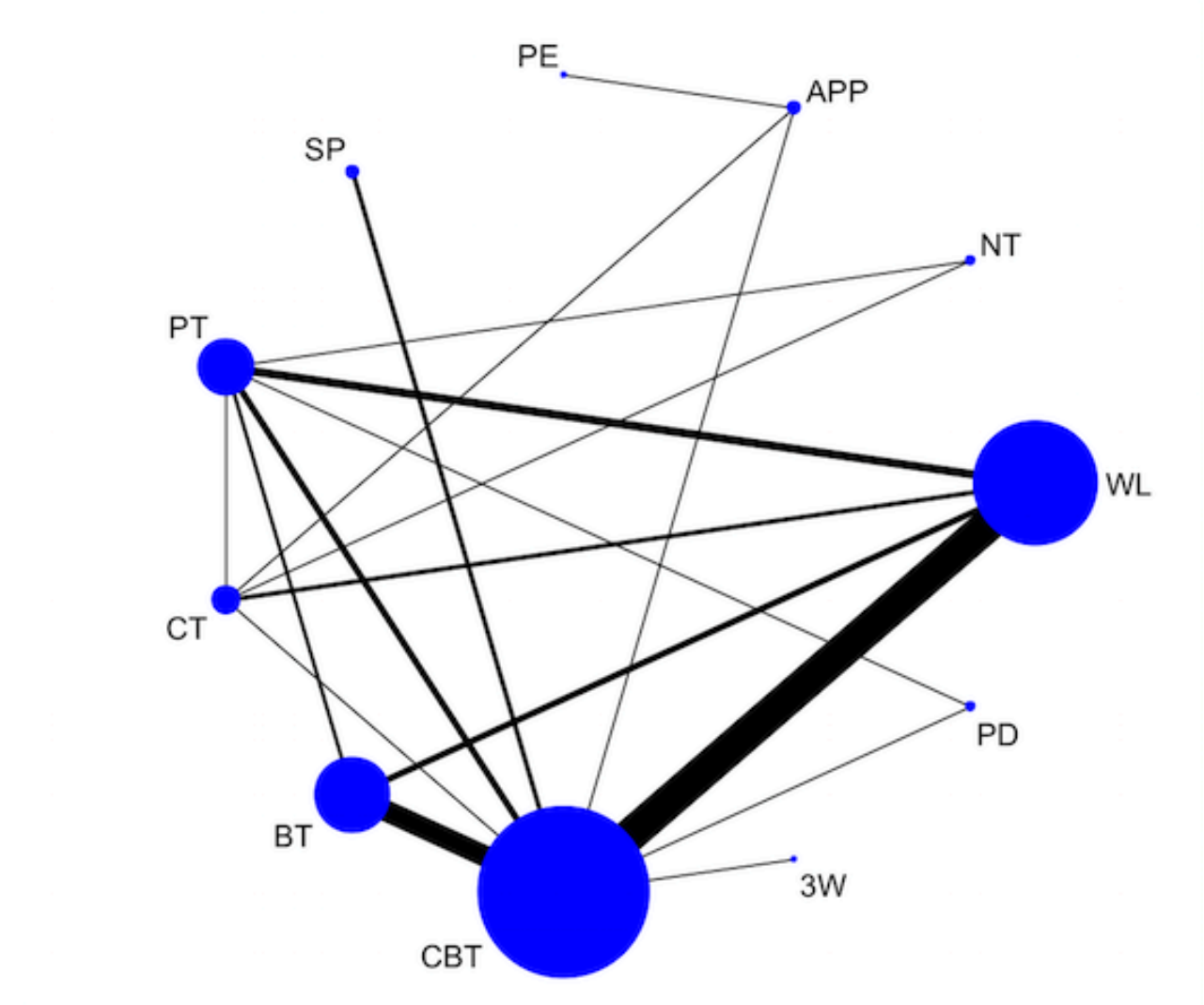
The ranking of treatments with respect to short-term response, according to the SUCRA value derived from NMA adjusted for small study effects, is presented in Table 6. We observed the highest rankings respectively for cognitive behaviour therapy (CBT), psychodynamic therapy (PD) and supportive psychotherapy (SP). Again, results regarding supportive psychotherapy must be interpreted with caution because SP is included in the network as a node with a single connection to the network, being compared only with CBT (three studies, OR 1.02, 95% CrI 0.38 to 2.73).

3. Dropouts for any reason in the short term

3.1 Network plot

Figure 18 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, ST-dropout data were available for all the psychological therapies and the control conditions considered for this review. CBT appeared to be the most studied intervention, followed by behaviour therapy (BT), physiological therapies (PT) and cognitive therapy (CT). Wait list (WL) was the most studied among comparator interventions. The most studied comparison was CBT versus WL. The network appeared to be moderately connected, with three interventions represented as nodes with a single connection to the network (SP, PE and 3W). Forty-seven studies including 2535 participants contributed data to this outcome.

Figure 18. Short-term dropouts: network plot



3.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

We excluded all studies with zero events in all arms (n=10) in the analyses. The study Ost 1993 is a three-arm study with zero events in two of the arms (CBT, PT) and one event in the third (BT). For the pairwise meta-analysis, we excluded the CBT versus PT comparison for this study and only kept the other two comparisons (CBT versus BT, PT versus BT).

As summarised in the left part of Table 7, direct evidence was available for 14 comparisons. For seven of these comparisons there was only one study available; for the remaining seven comparisons we performed a random-effects meta-analysis. As shown in the table, only two comparisons were informed by 10 or more studies, that is CBT versus WL (14 studies) and CBT versus BT (10 studies): their forest plots are respectively presented in Figure 19 and Figure 20. Only two comparisons showed statistically significant results: we found that WL was associated with significantly fewer dropouts than BT (four studies, OR 0.34, 95% CI 0.16 to 0.69), and that PD was associated with significantly fewer dropouts than PT (one study, OR 0.16, 95% CI 0.03 to 0.84).

Figure 19. Short-term dropouts: forest plot for the comparison WL vs CBT

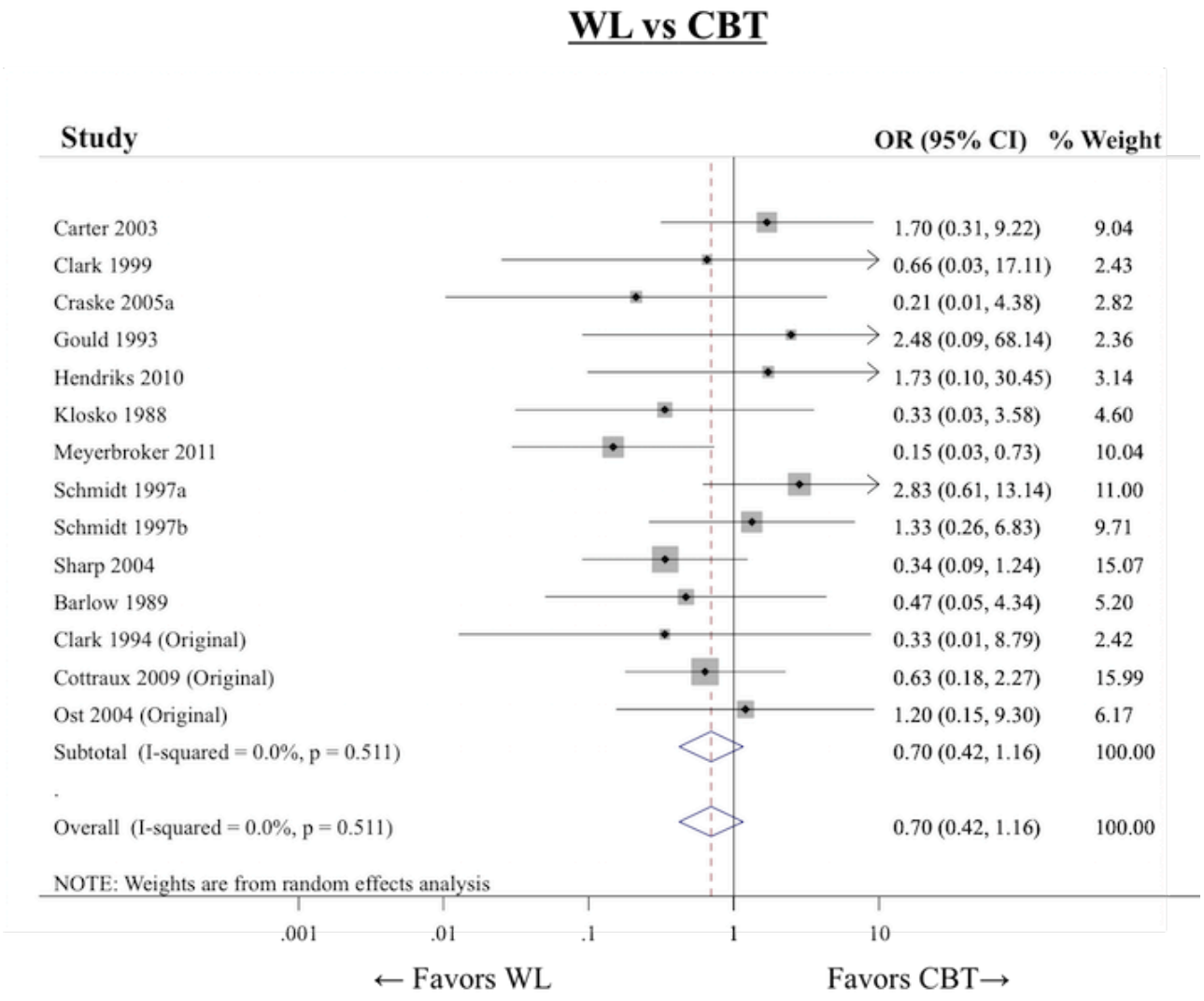
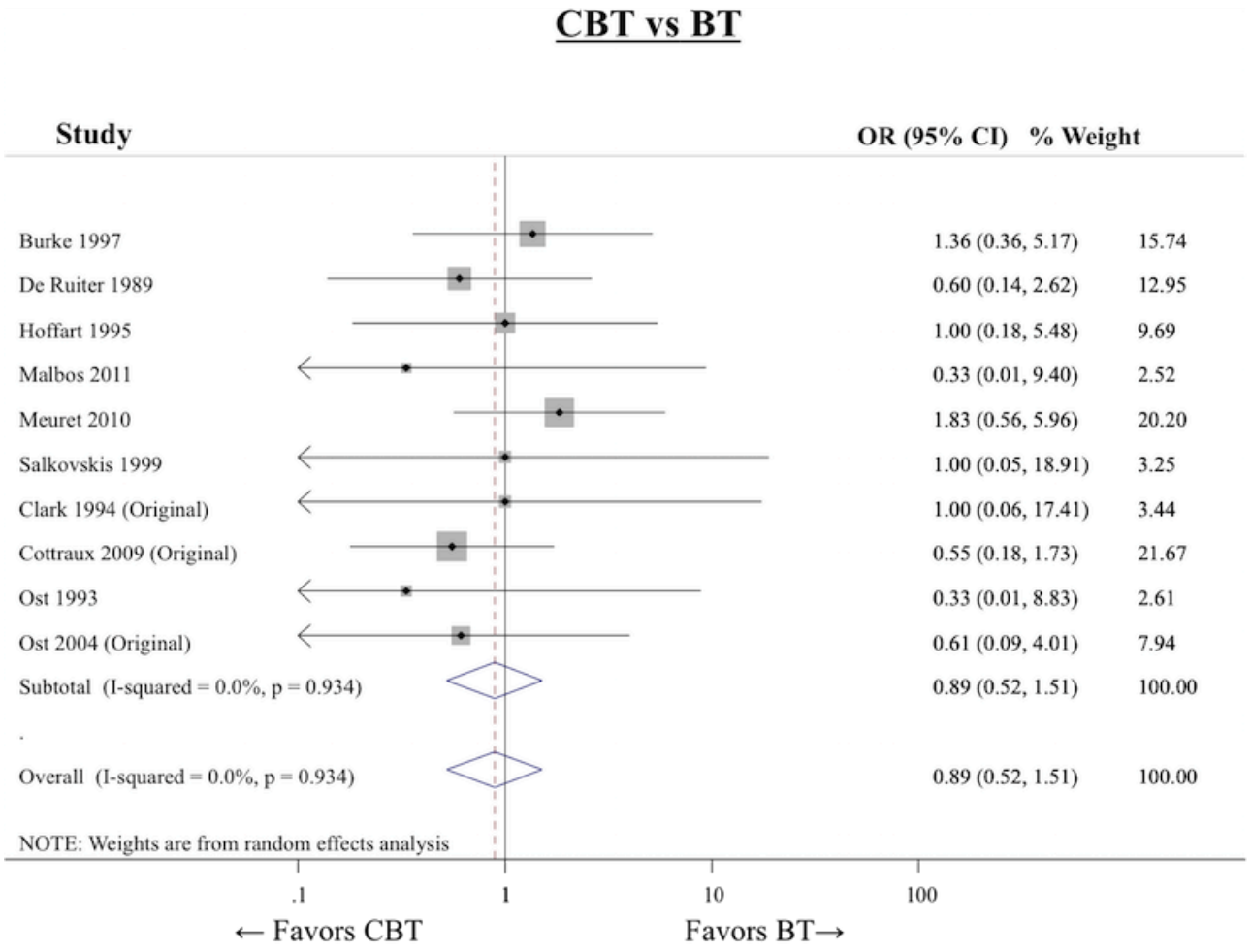


Figure 20. Short-term dropouts: forest plot for the comparison CBT vs BT



Heterogeneity

I^2 values and their 95% CIs, for the comparisons reported in three studies or more, are presented in Table 8. We observed the highest I^2 value in the comparisons CBT versus SP (I^2 = 49%). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) these values suggest that, in these two comparisons, a moderate percentage of the observed variability in the effect estimates was due to heterogeneity rather than sampling error (chance).

Small study effects

Following the protocol, we produced funnel plots for all comparisons appearing in more than 10 studies. There were two comparisons appearing in 10 studies or more, that is WL versus CBT (Figure 21) and CBT versus BT (Figure 22). We found no evidence of asymmetry in either of the funnel plots.

Figure 21. Short-term dropouts: funnel plot for the comparison WL vs CBT

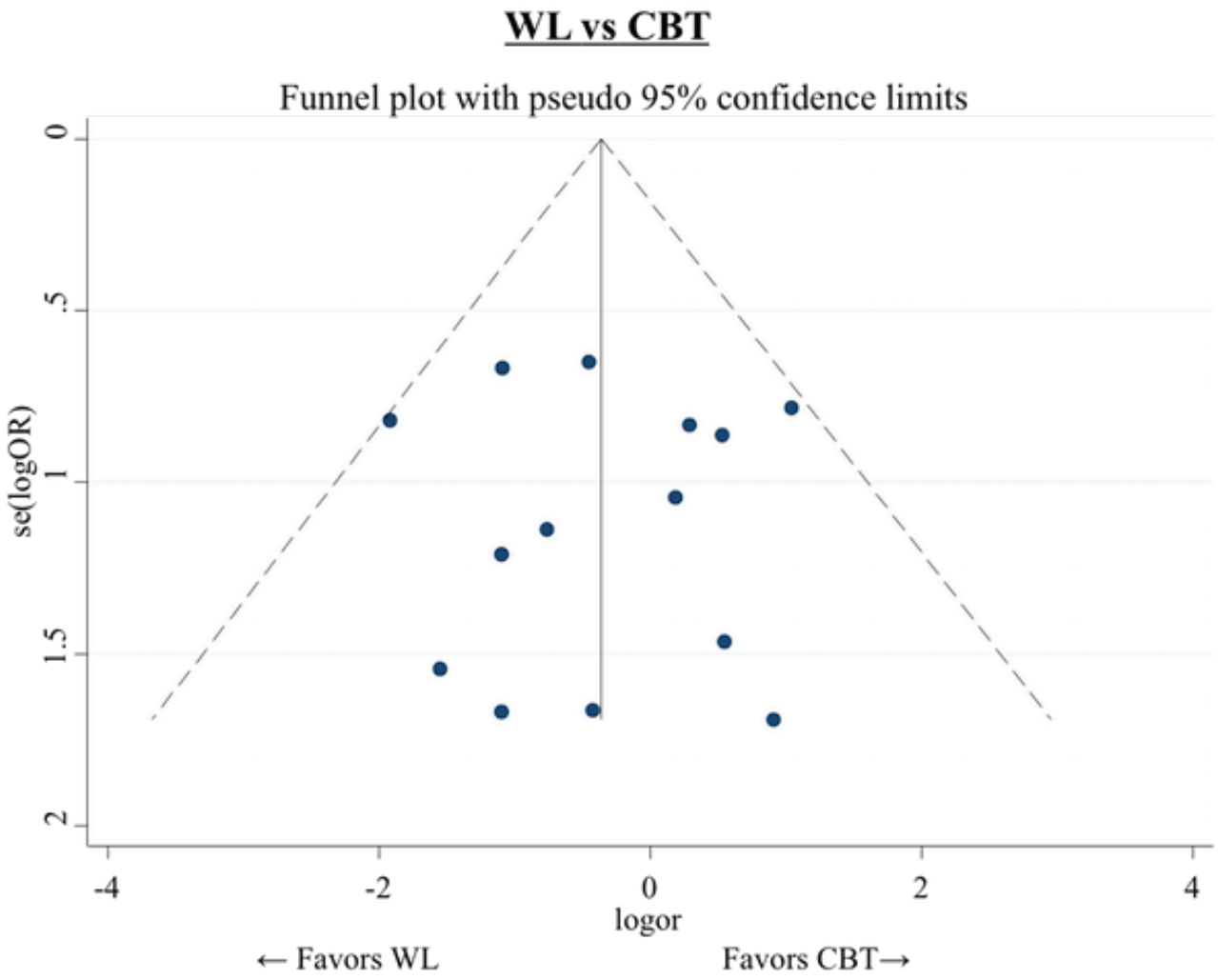
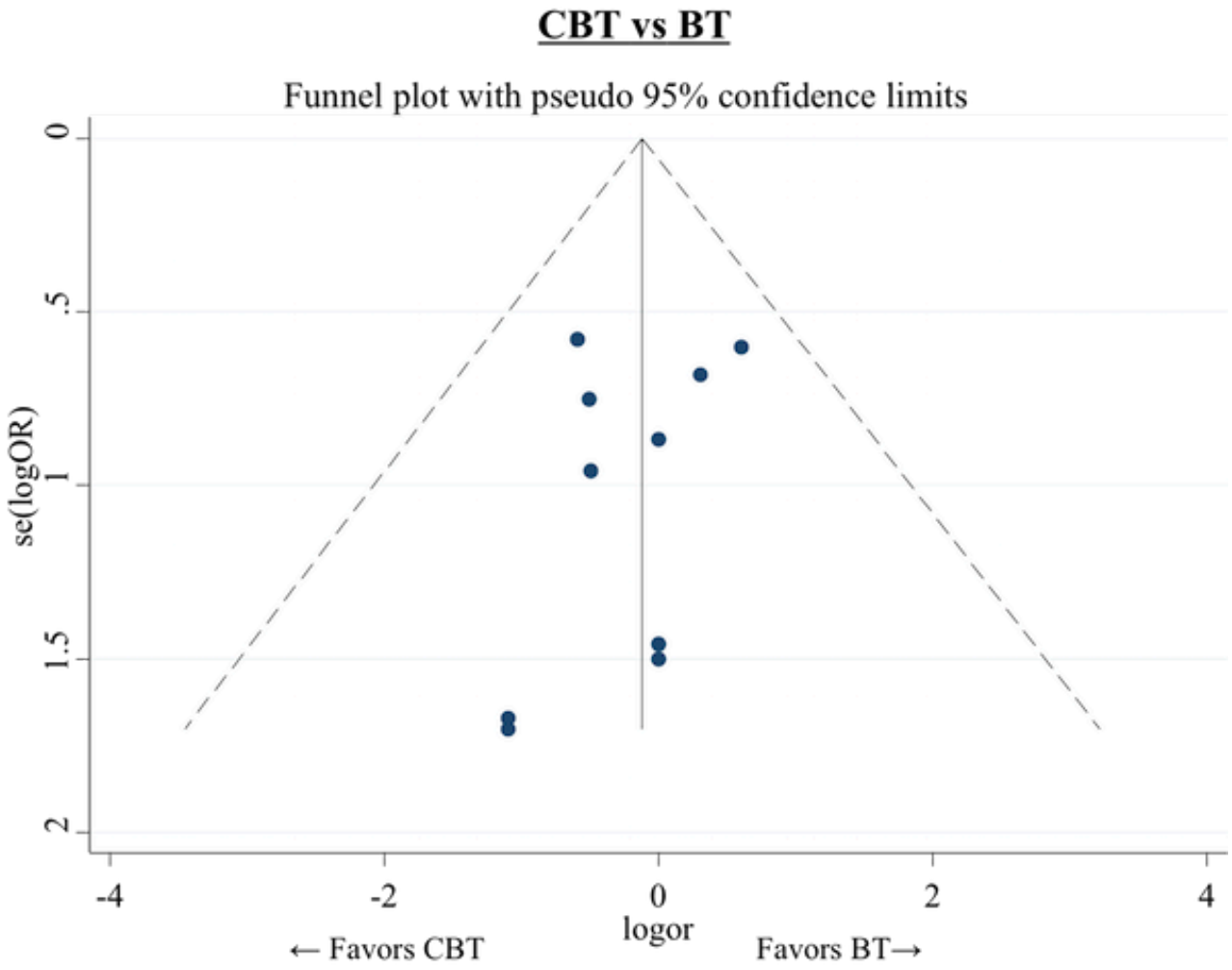


Figure 22. Short-term dropouts: funnel plot for the comparison CBT vs BT



3.3 Network meta-analysis and its inconsistency

Network meta-analysis

Results of the network meta-analysis (NMA) for short-term dropouts are presented in the right part of Table 7. Indirect evidence could be calculated for 41 comparisons for which direct evidence was unavailable. Results for comparisons WL versus BT remained statistically significant also within the context of NMA, with an OR of 0.52 (95% CI 0.30 to 0.93).

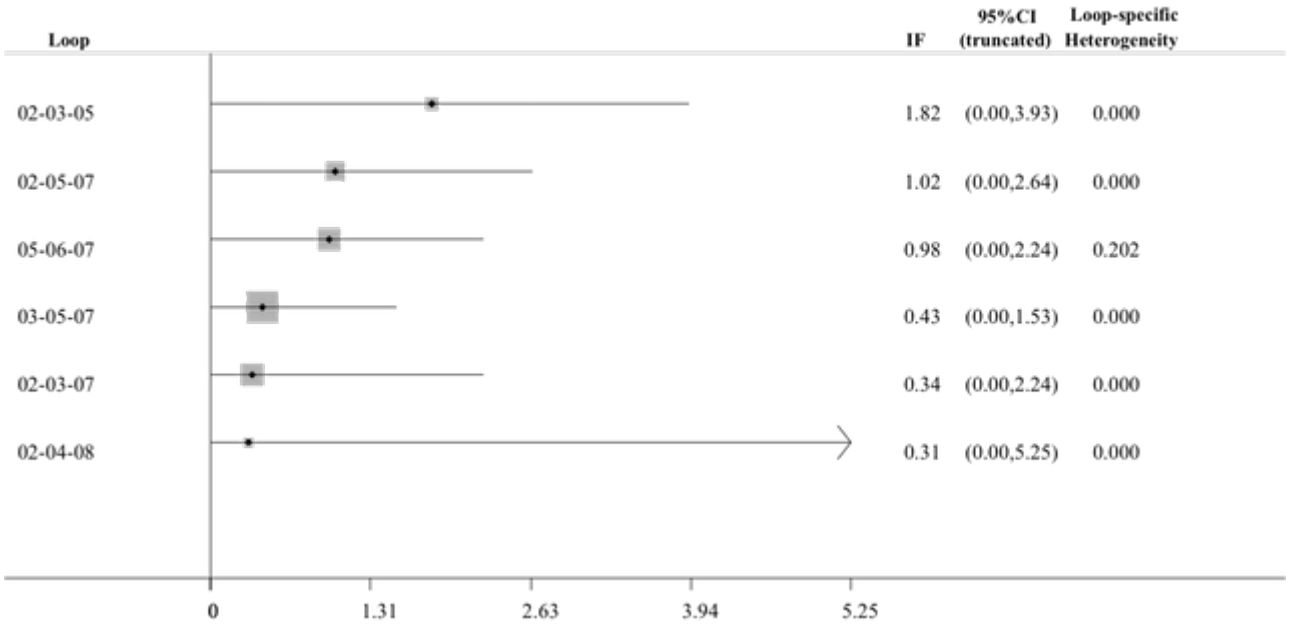
Network heterogeneity and inconsistency

The estimated value of heterogeneity lay well within the range of values usually found in Cochrane reviews, as presented by Turner 2012.

We compiled a table of important trial and patient characteristics including therapy duration and percentage of agoraphobic, depressed and drug treated patients. Its visual inspection showed that those effect modifiers were similarly distributed across comparisons in the network: we therefore concluded that there wasn't important evidence against the transitivity assumption.

Using a loop-specific approach, we found no evidence of inconsistency in the network. In Figure 23 we give all inconsistency factors for the network. The study Ost 1993 was a three-arm study with zero events in two of the arms (CBT, PT) and one event in the third (BT). For estimating inconsistency we excluded the CBT-PT comparison from this study. As a sensitivity analysis we excluded the study Ost 1993 but there were no qualitative changes in the inconsistency factors.

Figure 23. Short-term dropouts: inconsistency factors for the network



The test for global inconsistency (design-by-treatment inconsistency) provided no proof of inconsistency in the network (Chi^2 = 11.67 with 14 degrees of freedom; P value = 0.63).

3.4 Ranking of treatments

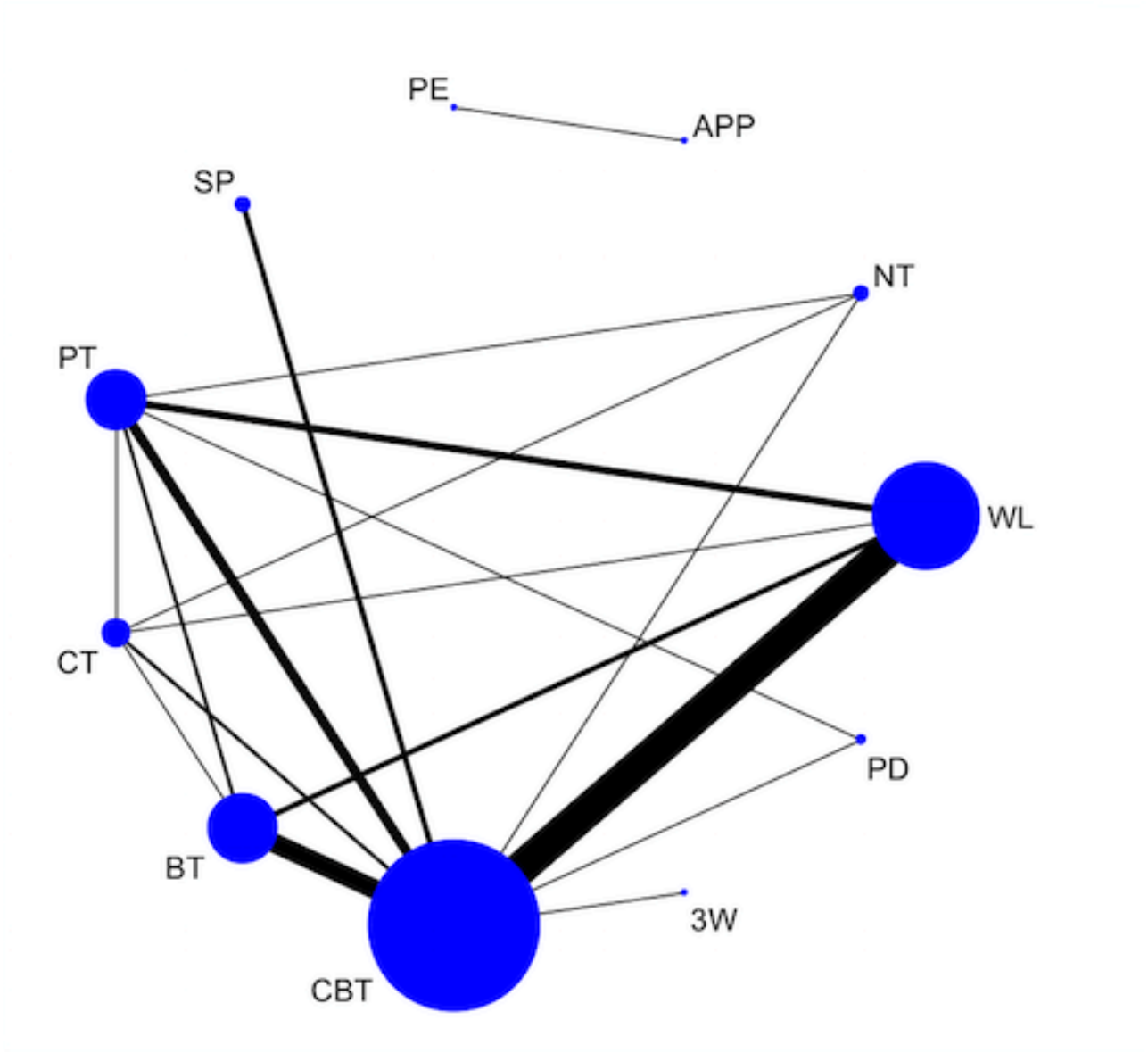
The ranking of treatments with respect to short-term dropouts, according to the SUCRA value derived from NMA, is presented in Table 9. We observed the highest rankings (that correspond with a lower dropout rate) respectively for no treatment (NT), psychodynamic therapy (PD) and third-wave CBT (3W).

4. Short-term improvement of panic disorder with or without agoraphobia as measured on a continuous scale

4.1 Network plot

Figure 24 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, ST-improvement data were available for all the psychological therapies and the control conditions considered for this review. CBT appeared to be the most studied intervention, followed by behaviour therapy (BT), physiological therapies (PT) and cognitive therapy (CT). Wait list (WL) was the most studied among comparator interventions. The most studied comparison was CBT versus WL. The network appeared to be poorly connected, with two interventions represented as nodes with a single connection (SP and 3W); two interventions were connected only with each other but not with the rest of the network (PE and APP), so they could not be included in the analyses. Excluding those four interventions, the rest of the network appeared to be moderately well connected. Fifty-seven studies including 2318 participants contributed data to this outcome.

Figure 24. Short-term improvement: network plot



4.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As summarised in the left part of Table 10, direct evidence was available for 18 comparisons. For 11 of these comparisons there was only one study available; for the remaining seven comparisons we performed a random-effects meta-analysis. As shown in the table, only two comparisons were informed by 10 or more studies, that is CBT versus WL (17 studies) and CBT versus BT (10 studies): their forest plots are respectively presented in Figure 25 and Figure 26. Among psychological therapies, three were shown to be significantly better than WL in terms of short-term improvement: PT (four studies: standardised mean difference (SMD) 0.87, 95% CI 0.09 to 1.65), BT (three studies: SMD 0.92, 95% CI 0.59 to 1.26) and CBT (seventeen studies: SMD 1.14, 95% CI 0.87 to 1.41). CBT was also found to be significantly better than NT (one study: SMD 1.30, 95% CI 0.46 to 2.14). Finally, two comparisons among two active treatments showed a statistically significant difference in terms of short-term improvement: CBT versus BT (10 studies: SMD -0.24 in favour of CBT, 95% CI -0.45 to -0.03) and PD versus PT (one study: SMD -1.18 in favour of PD, 95% CI -1.59 to -0.57).

Figure 25. Short-term improvement: forest plot for the comparison WL vs CBT

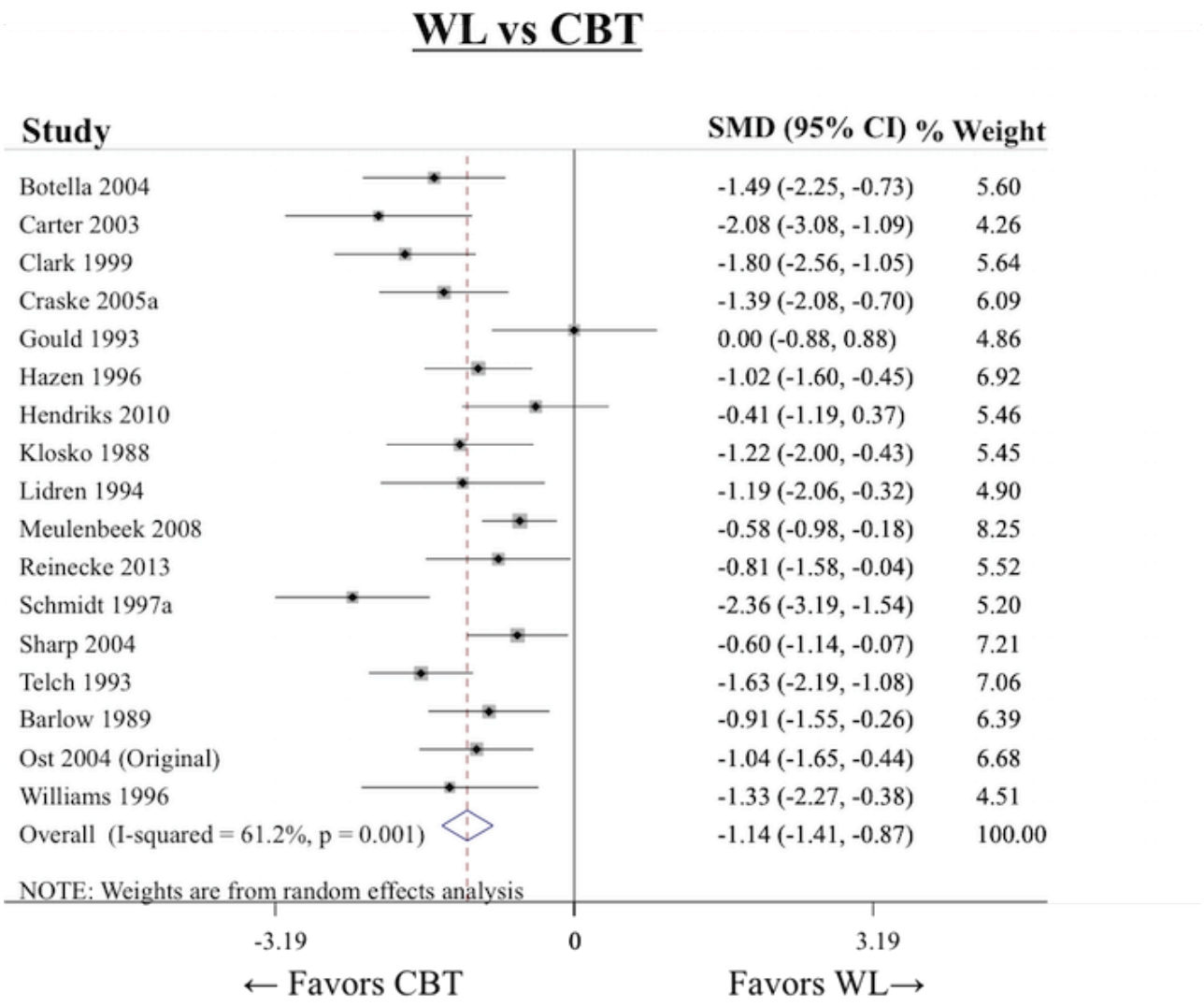
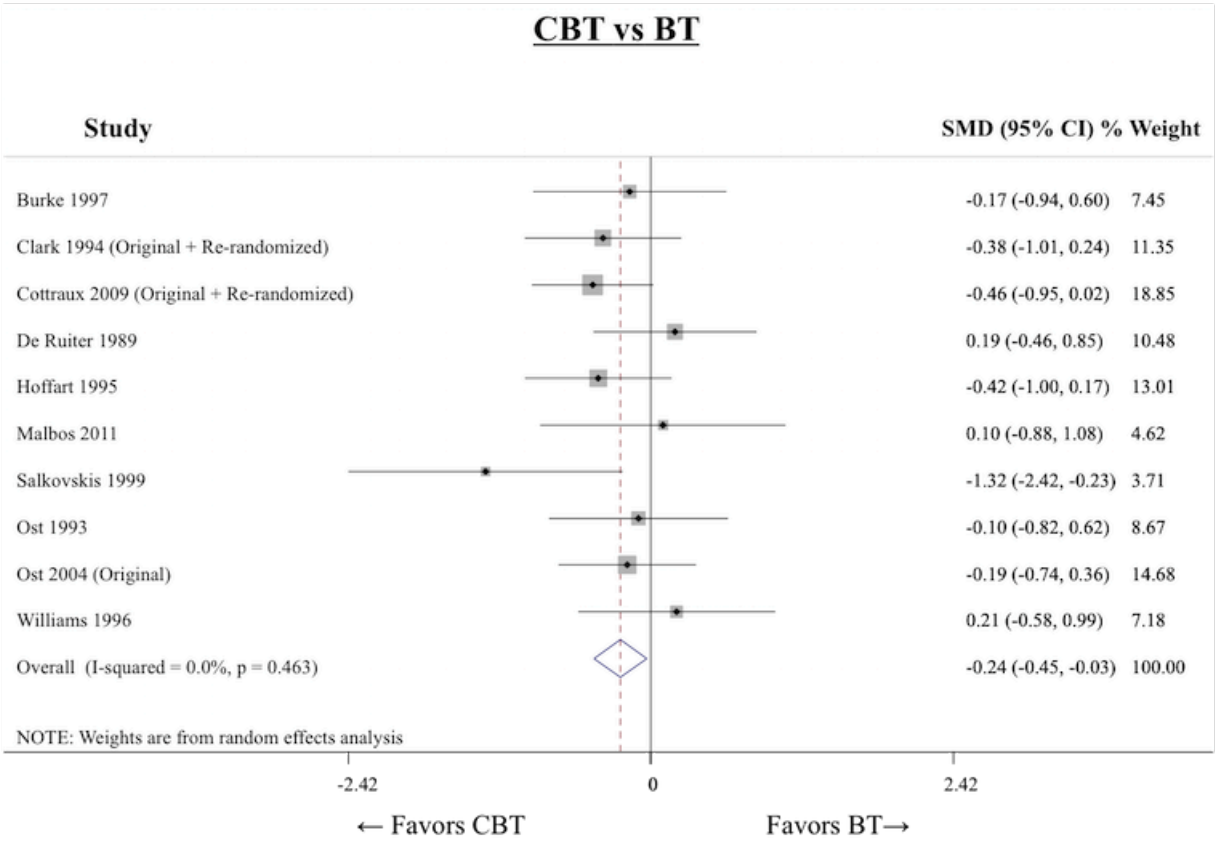


Figure 26. Short-term improvement: forest plot for the comparison CBT vs BT



Heterogeneity

I^2 values and their 95% CIs, for the comparisons reported in three studies or more, are presented in Table 11. As shown in the table, we observed the highest I^2 values in the comparisons WL versus PT (I^2 = 79%) and WL versus CBT (I^2 = 61%). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) these values suggest that, in these two comparisons, a substantial percentage of the observed variability in the effect estimates was due to heterogeneity rather than sampling error (chance). In the comparison WL versus PT (four studies), heterogeneity appeared to be related to the study Meuret 2008, which shows an unexpectedly high SMD in favour of PT, possibly related to a strong researcher allegiance bias (see Characteristics of included studies); in the case of WL versus CBT (17 studies), heterogeneity seems to be mainly due to two outlier studies, one showing a very high SMD in favour of CBT (Schmidt 1997a, SMD -2.36, 95% CI -3.19 to -1.54) and the other showing no effect in either direction (Gould 1993, SMD 0.00, 95% CI -0.88 to 0.88).

Small study effects

Following the protocol, we produced funnel plots for all comparisons appearing in more than 10 studies. There were two comparisons appearing in 10 studies or more, that is CBT versus WL (Figure 27) and CBT versus BT (Figure 28). We found no evidence of asymmetry in either of the funnel plots.

Figure 27. Short-term improvement: funnel plot for the comparison WL vs CBT

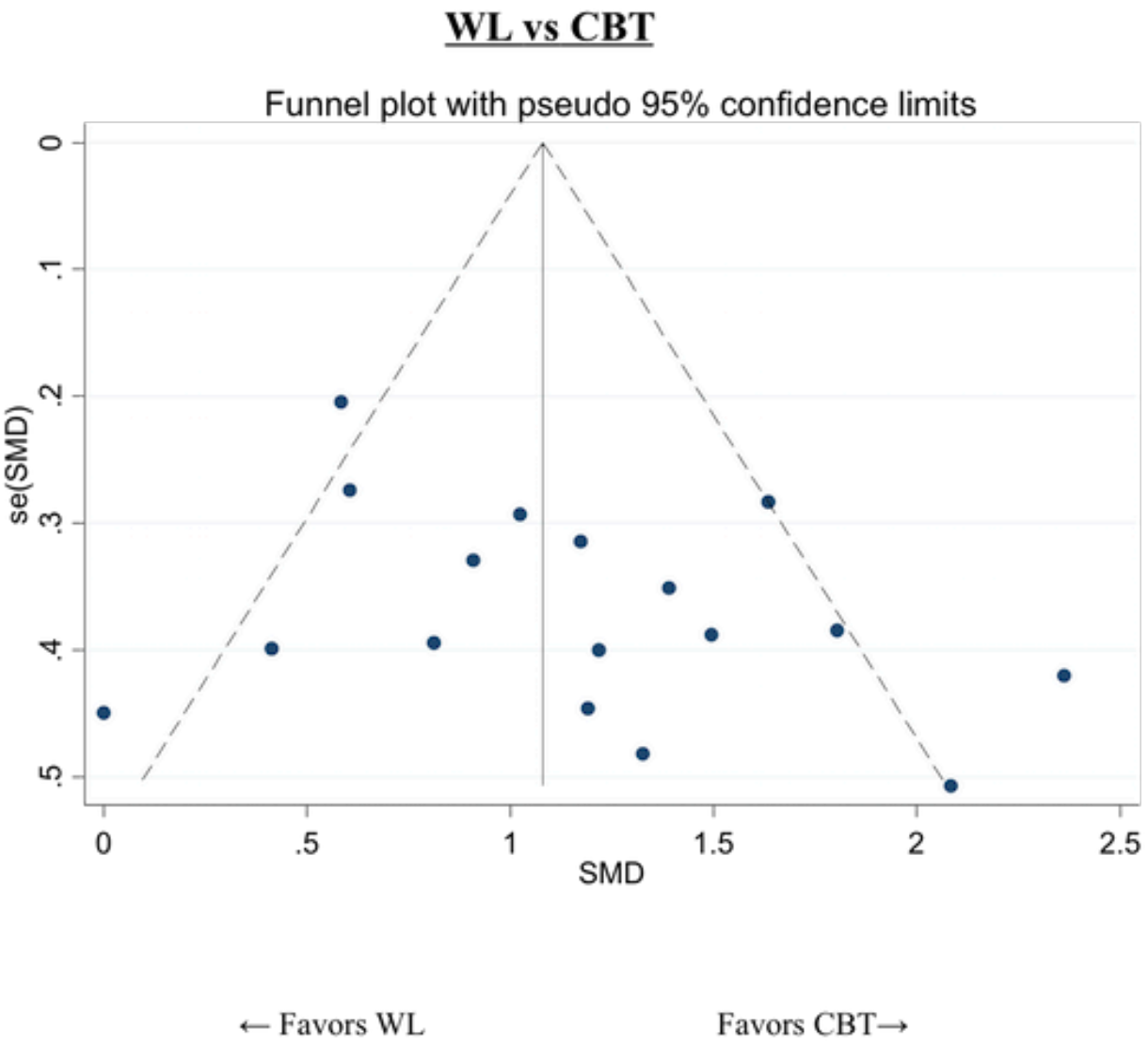
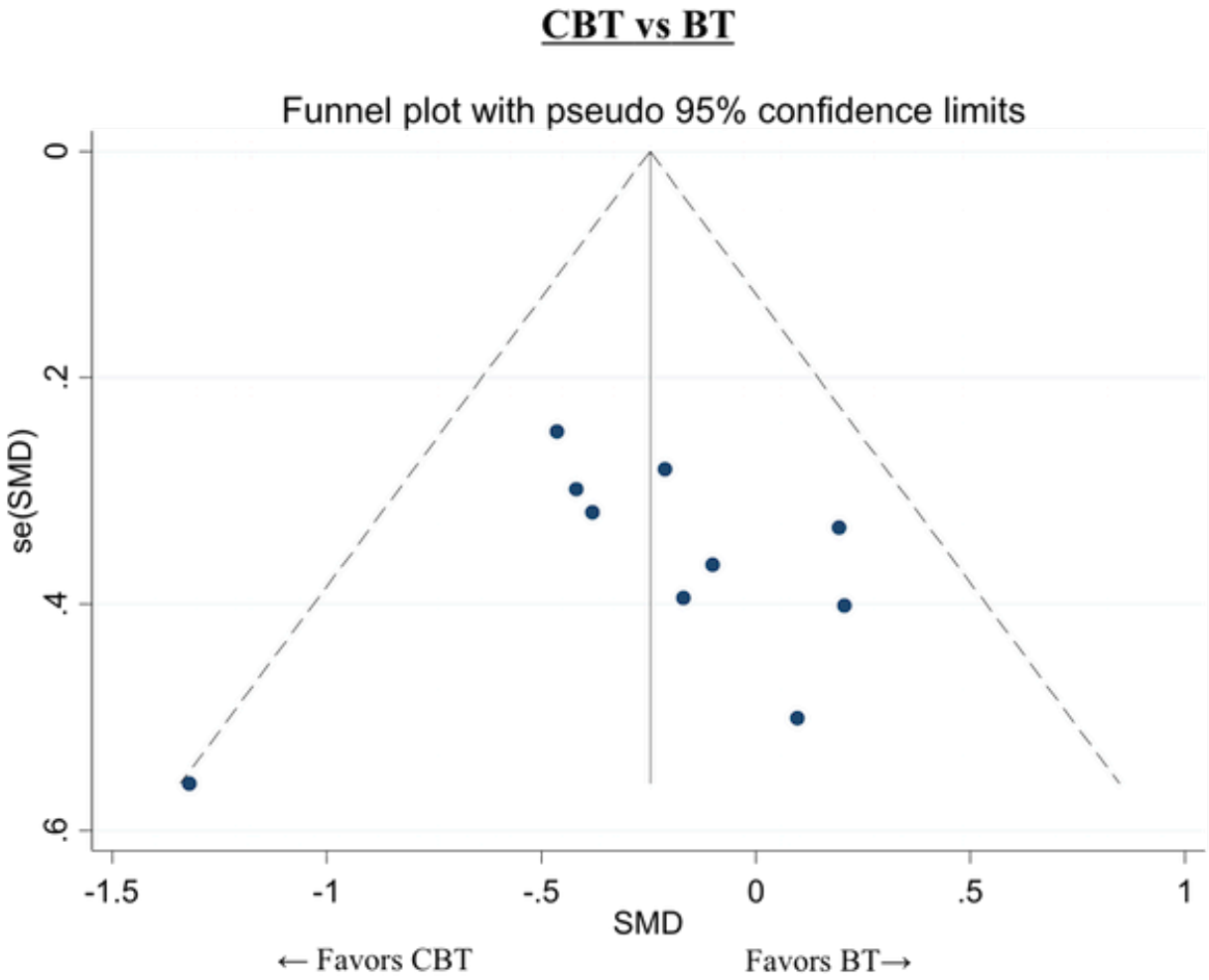


Figure 28. Short-term improvement: funnel plot for the comparison CBT vs BT



4.3 Network meta-analysis and its inconsistency

Network meta-analysis

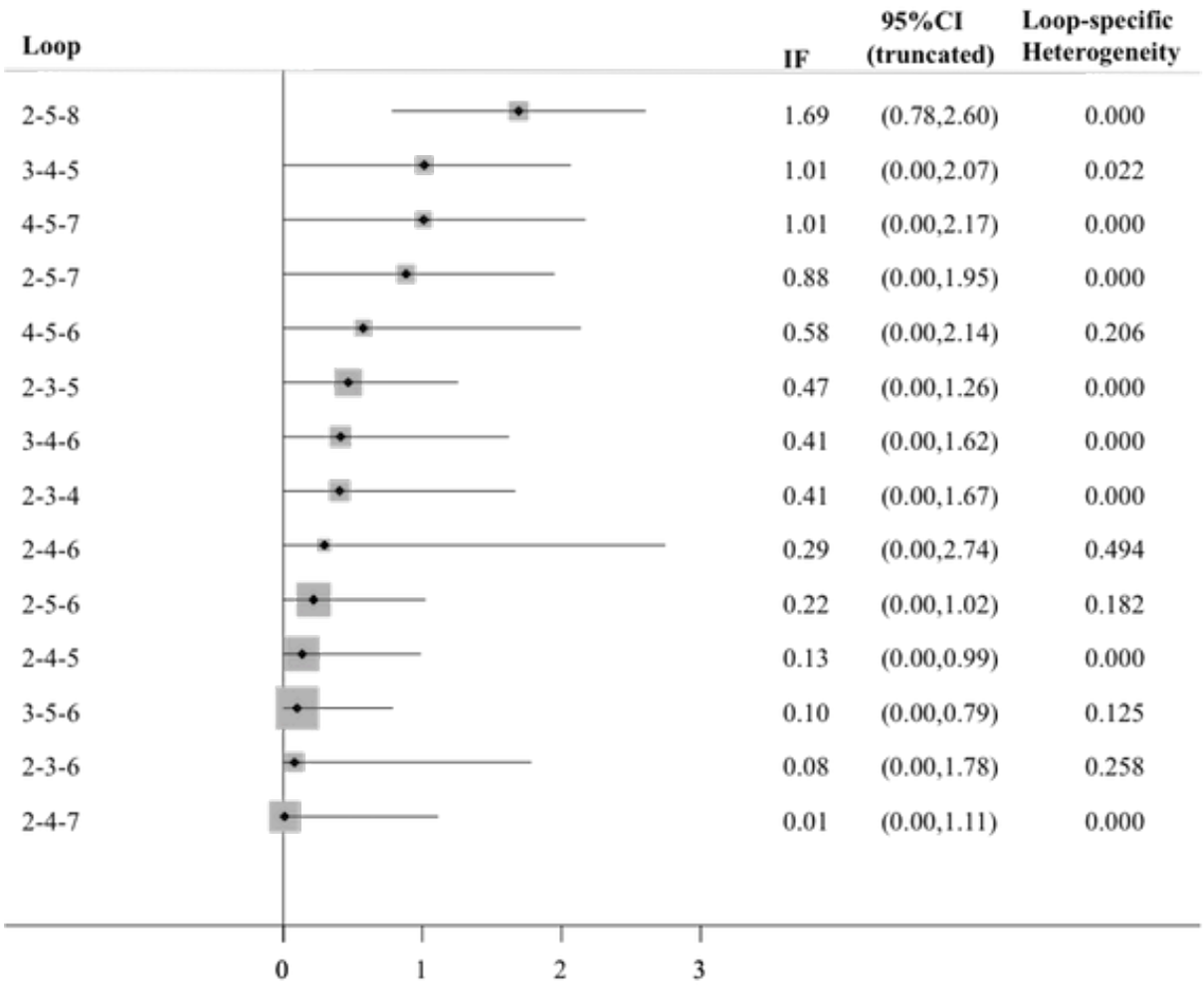
Results of the network meta-analysis (NMA), for short-term improvement on a continuous scale, are presented in the right part of Table 10. Indirect evidence could be calculated for 19 comparisons for which direct evidence was unavailable. The comparison PE versus APP, for which direct evidence was available, is not included in the NMA because it is disconnected from the rest of the network (see Figure 24). Four comparisons remained statistically significant also within the context of NMA: WL versus PT (SMD 0.80, 95% CI 0.47 to 1.13), WL versus BT (SMD 0.89, 95% CI 0.57 to 1.20), WL versus CBT (SMD 1.09, 95% CI 0.88 to 1.31) and NT versus CBT (SMD 0.83, 95% CI 0.16 to 1.50). The comparisons CBT versus BT and PD versus PT lost significance in the NMA. We also found supportive psychotherapy (SP) and cognitive therapy (CT) to be significantly better than WL, showing a SMD of respectively 1.05 (95% CI 0.49 to 1.60) and 0.88 (95% CI 0.34 to 1.42).

Inconsistency

We compiled a table of important trial and patient characteristics including therapy duration and percentage of agoraphobic, depressed and drug treated patients. Its visual inspection showed that those effect modifiers were similarly distributed across comparisons in the network: we therefore concluded that there wasn't important evidence against the transitivity assumption.

Using a loop-specific heterogeneity (where we allowed the same τ for all comparisons in a loop), we identified one inconsistent loop out of a total of 14 loops. The inconsistent loop was PT-CBT-PD (IF 1.79, 95% CI 0.78 to 2.60; see Figure 29 for other inconsistency factors for the network). This loop was found to be inconsistent also when allowing for a comparison-specific heterogeneity. Although in NMA 5% of the loops are expected to be inconsistent by chance, it must be noted that this same loop showed the highest inconsistency factor (although non-significant) also in the NMA for ST-remission and ST-response. Furthermore, two of the three edges of the loop were only informed by one study each, both considered to be at high risk of researcher allegiance bias. We can summarise the inconsistency in available direct evidence as follows: CBT appeared to perform better than PD (one study: SMD 0.57, 95% CI -0.07 to 1.20) and PD appeared to perform better than PT (one study: SMD -1.18, 95% CI -1.79 to -0.57); however, the comparison CBT versus PT showed almost no difference between the two treatments (five studies: SMD -0.05, 95% CI -0.3 to 0.19).

Figure 29. Short-term improvement: inconsistency factors for the network



The global test for inconsistency did not reveal any definite proof about inconsistency in the network (Chi^2 = 9.13 with 14 degrees of freedom, P value 0.82).

4.4 Ranking of treatments

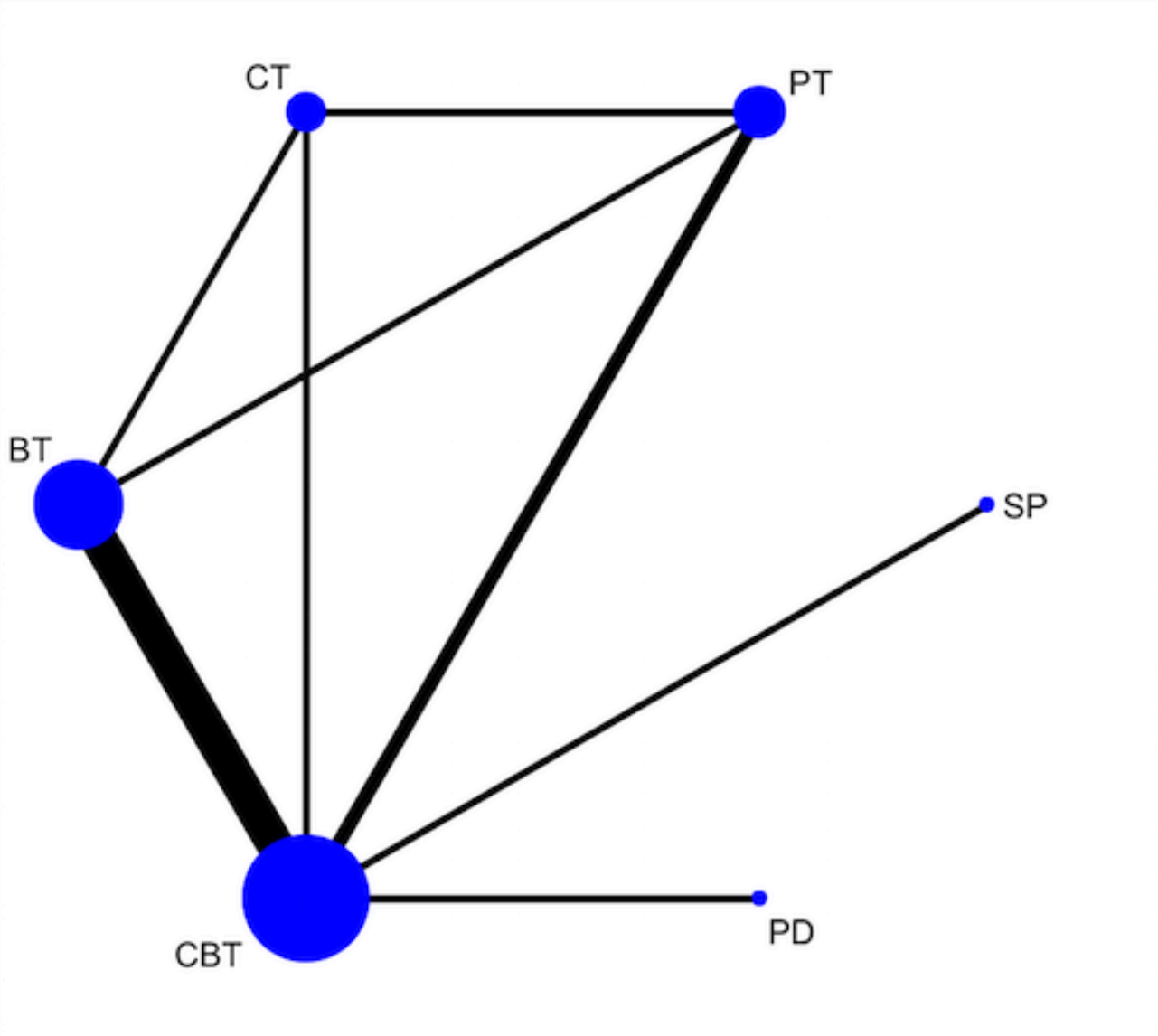
The ranking of treatments with respect to short-term improvement, according to the SUCRA value derived from NMA, is presented in Table 12. We observed the highest rankings respectively for psychodynamic psychotherapy (PD), cognitive behaviour therapy (CBT) and supportive psychotherapy (SP). However, results regarding PD must be interpreted with caution because they relied on only two studies (Beutel 2013; Milrod 2006a), both included in the inconsistent loop PT-CBT-PD mentioned above.

5. Long-term remission or response of panic disorder with or without agoraphobia

5.1 Network plot

Figure 30 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, long-term data were available for only six active interventions. No study explored long-term (LT)-remission/response for third-wave CBT (3W) and psychoeducation (PE), nor for any control condition (WL, NT, APP). CBT was the most studied intervention, followed by behaviour therapy (BT), physiological therapies (PT) and cognitive therapy (CT). The most studied comparison was CBT versus BT. With the exception of SP and PD, both represented as nodes compared only with CBT, the network appeared to be well connected. Nine studies including 464 participants contributed data to this outcome.

Figure 30. Long-term remission/response: network plot



5.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As summarised in the left part of Table 13, direct evidence was available for eight comparisons. For six of these comparisons there was only one study available; for the remaining two comparisons we performed a random-effects meta-analysis. As shown in the table, no comparison was informed by 10 or more studies. None of the available comparisons showed statistically significant differences, in terms of long-term remission/response, between active treatments. However, this finding must be interpreted while taking into account the exiguity of available long-term data.

Heterogeneity

Only the comparison CBT versus BT was informed by more than two studies and it showed an I^2 value of 37%, with a 95% CI going from 0% to 77% (standard pairwise meta-analysis). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) this value suggests that the percentage of the observed variability in the effect estimates due to heterogeneity rather than sampling error may not have been important.

Small study effects

For this outcome, no comparison was informed by enough studies to produce a funnel plot in order to explore the presence of small study effects.

5.3 Network meta-analysis and its inconsistency

Network meta-analysis

Results of the network meta-analysis (NMA) for long-term remission/response are presented in the right part of Table 13. Indirect evidence could be calculated for seven comparisons for which direct evidence was unavailable. None of the available comparisons showed statistically significant differences between active treatments. Again, this finding must be interpreted while taking into account the exiguity of available long-term data.

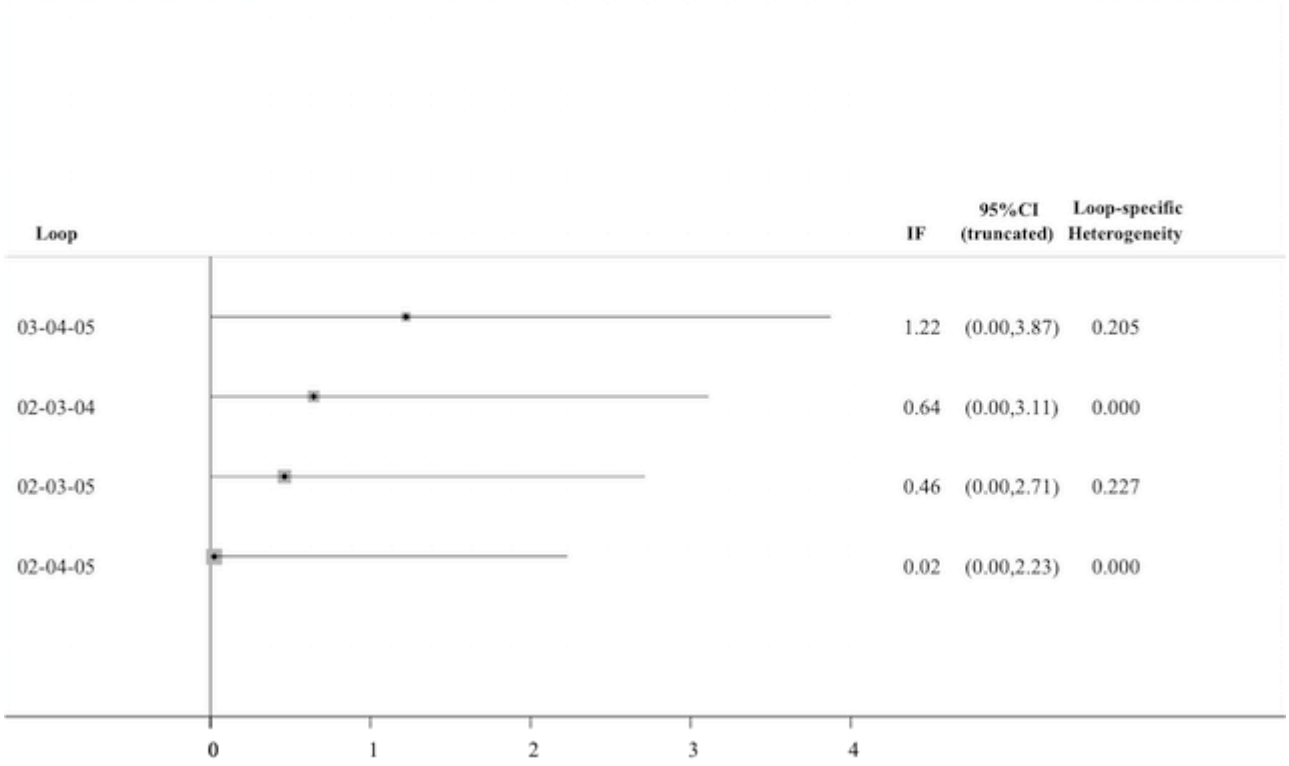
Network heterogeneity and inconsistency

The estimated value of heterogeneity lay well within the range of values usually found in Cochrane reviews, as presented by Turner 2012.

We compiled a table of important trial and patient characteristics including therapy duration and percentage of agoraphobic, depressed and drug treated patients. Its visual inspection showed that those effect modifiers were similarly distributed across comparisons in the network: we therefore concluded that there wasn't important evidence against the transitivity assumption.

Using a loop-specific heterogeneity (where we allow the same τ for all comparisons in a loop), we found no evidence of inconsistency in the network. In Figure 31 we give all inconsistency factors for the network.

Figure 31. Long-term remission/response: inconsistency factors for the network



The global test for inconsistency showed no signs of inconsistency in the network (Chi^2 = 1.80 with 4 degrees of freedom, P value = 0.77).

5.4 Ranking of treatments

The ranking of treatments with respect to long-term remission/response, according to the SUCRA value derived from NMA, is presented in Table 14. The highest ranking was observed for cognitive behaviour therapy, followed by psychodynamic therapy.

6. Subgroup analyses

The following analyses were aimed at exploring possible sources of heterogeneity for the first of the primary outcomes, that is short-term remission of panic disorder with or without agoraphobia.

6.1 Year of publication

For the two pairwise comparisons for which there were enough studies available (CBT versus BT and CBT versus WL) we performed a meta-regression to investigate the effect of the year of publication. In both pairwise comparisons older studies seemed to favour CBT (compared to BT and WL) but the effect was not found to be statistically significant in either of the cases. It must be noted that older studies tended to be smaller than newer ones. Thus, the small trend found was probably due to SSE. The meta-regression parameters were, respectively, -0.02 (95% CI -0.10 to 0.05) for the comparison CBT versus BT, and 0.04 (95% CI -0.06 to 0.15) for the comparison WL versus CBT.

6.2 Mean number of treatment sessions

Following the protocol, we divided the included studies that reported the mean number of treatment sessions into three groups. The first group was for studies with fewer than four sessions, the second was for studies with four to 12 sessions and the last group was for studies with more than 12 sessions. We performed a subgroup analysis for comparisons reported by enough studies (CBT versus BT and CBT versus WL).

Figure 32 shows the forest plot for the comparison WL versus CBT, subgrouping the studies that reported the mean number of treatment sessions, along with their meta-analysis (note that Clark 1999 had two CBT arms, one with 15 sessions and one with eight: for the purposes of this analysis we broke this study in two). It is evident from the plot that the number of sessions was not associated with the treatment effect (test for subgroup differences: P value = 0.937; I^2 = 0.0%). A meta-regression also showed no difference between

the groups and no reduction in heterogeneity. A meta-regression on the exact number of treatment sessions also showed no effect (Figure 33).

Figure 32. Subgroup analysis: number of treatment sessions, forest plot for the comparison WL vs CBT

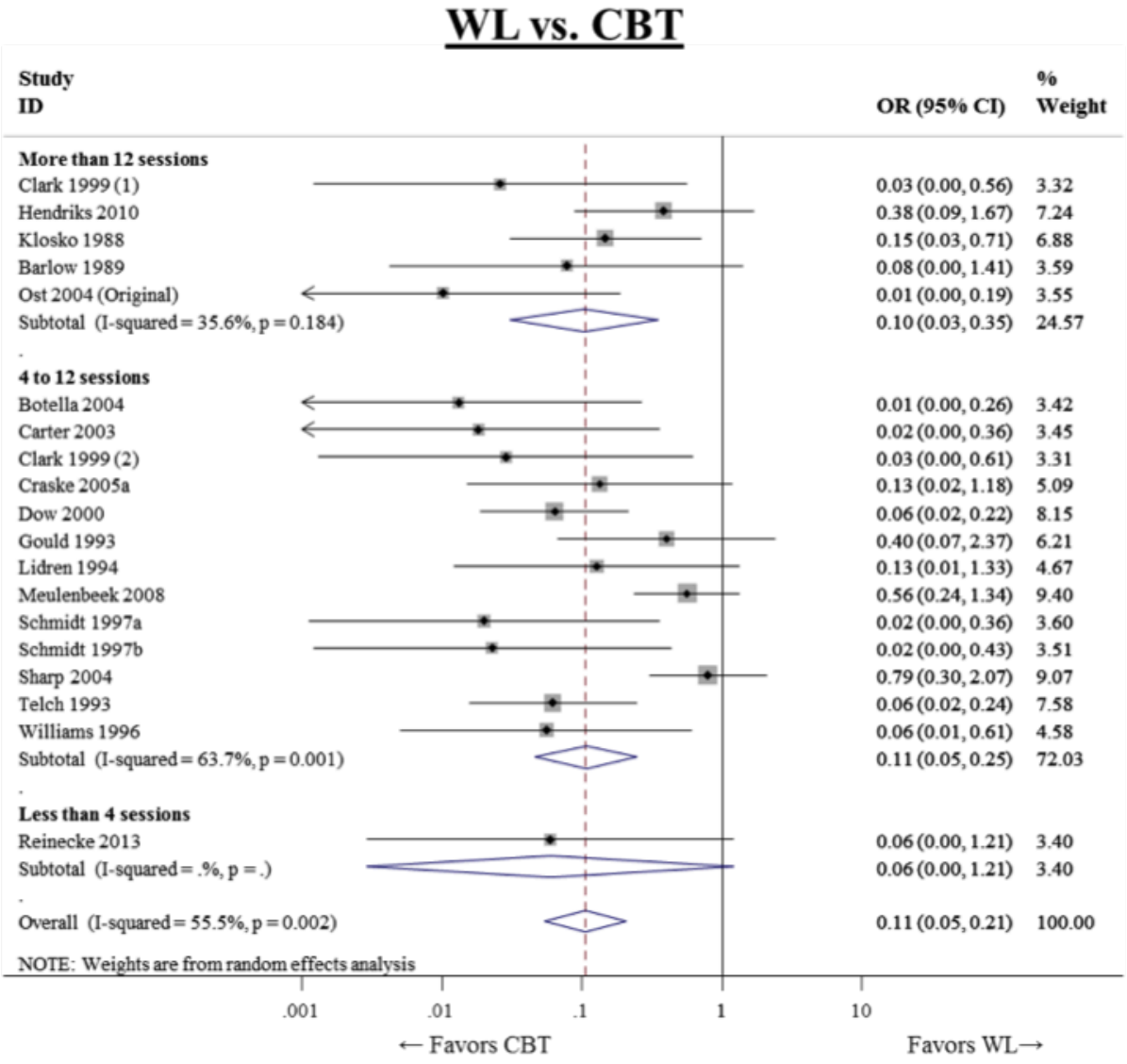


Figure 33. Subgroup analysis: number of treatment sessions, regression line for the comparison WL vs CBT

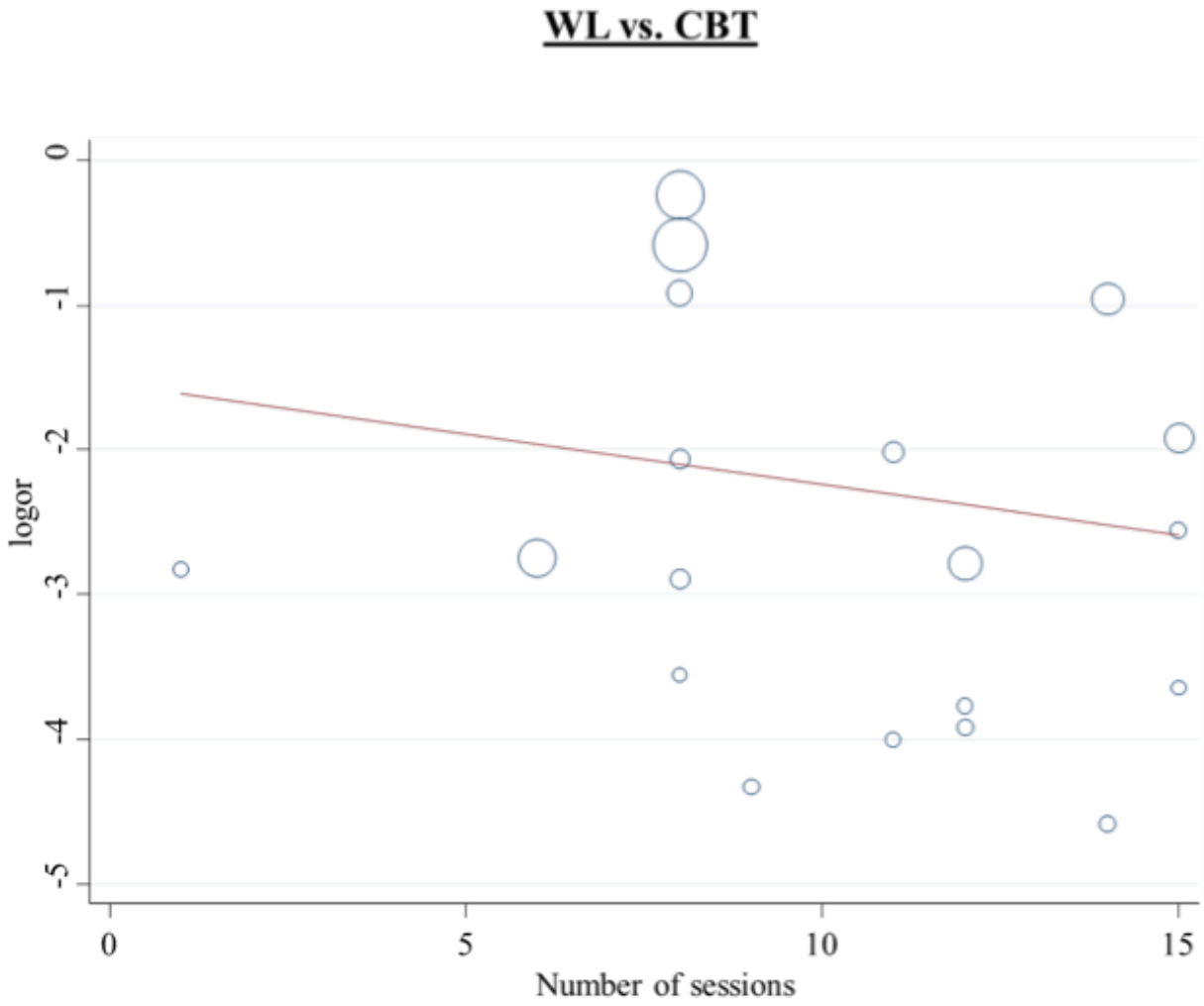


Figure 34 shows the forest plot for the comparison CBT versus BT, subgrouping the studies that reported the mean number of treatment sessions, along with their meta-analysis. Again, the plot showed no important differences among the two available groups (test for subgroup differences: P value = 0.269; I^2 = 18.2%). Results of the meta-regression on the (standardised) number of sessions showed that the increase in the number of sessions tended to increase the relative efficacy of CBT versus BT (Figure 35), but the increase was not statistically significant.

Figure 34. Subgroup analysis: number of treatment sessions, forest plot for the comparison CBT vs BT

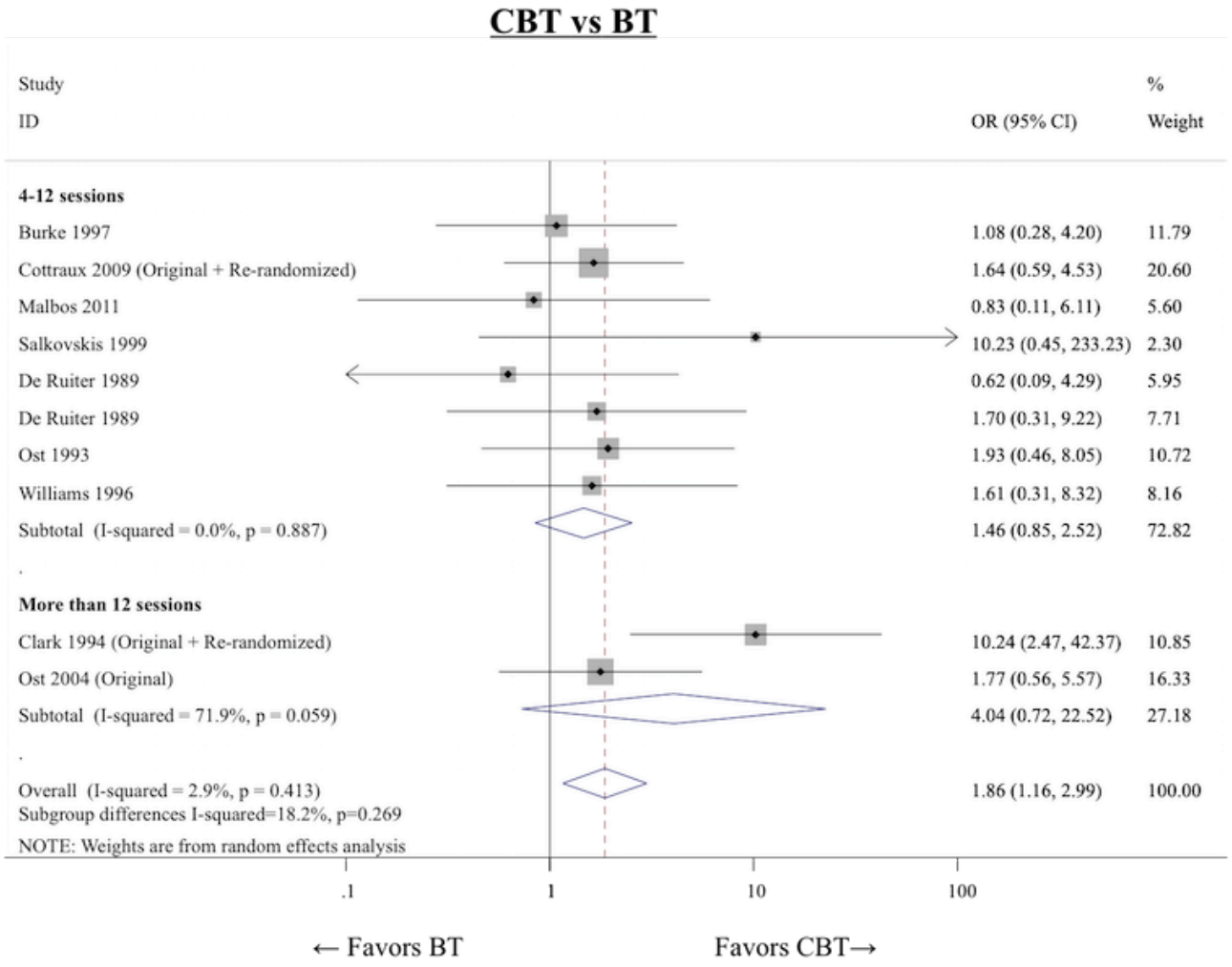
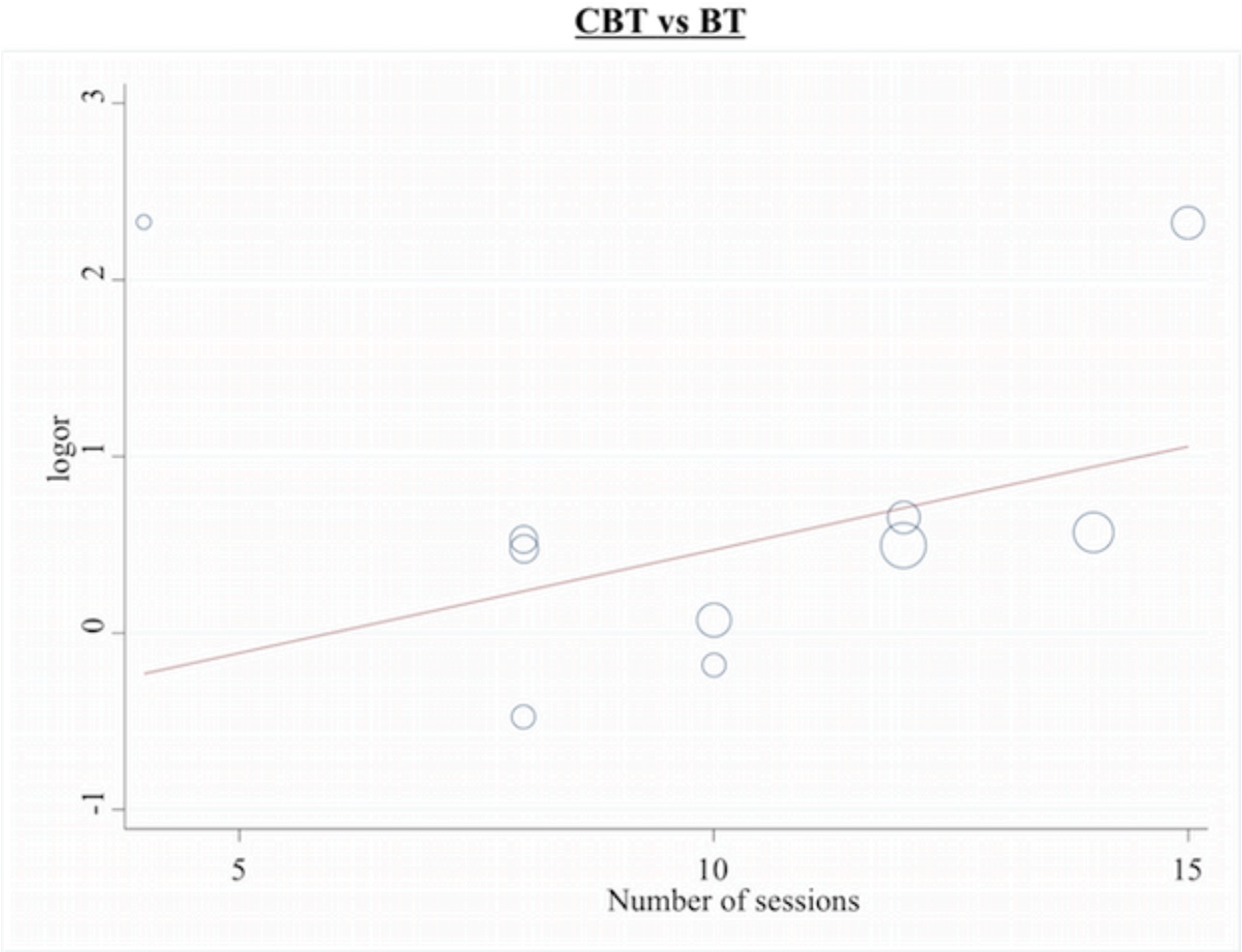


Figure 35. Subgroup analysis: number of treatment sessions, regression line for the comparison CBT vs BT

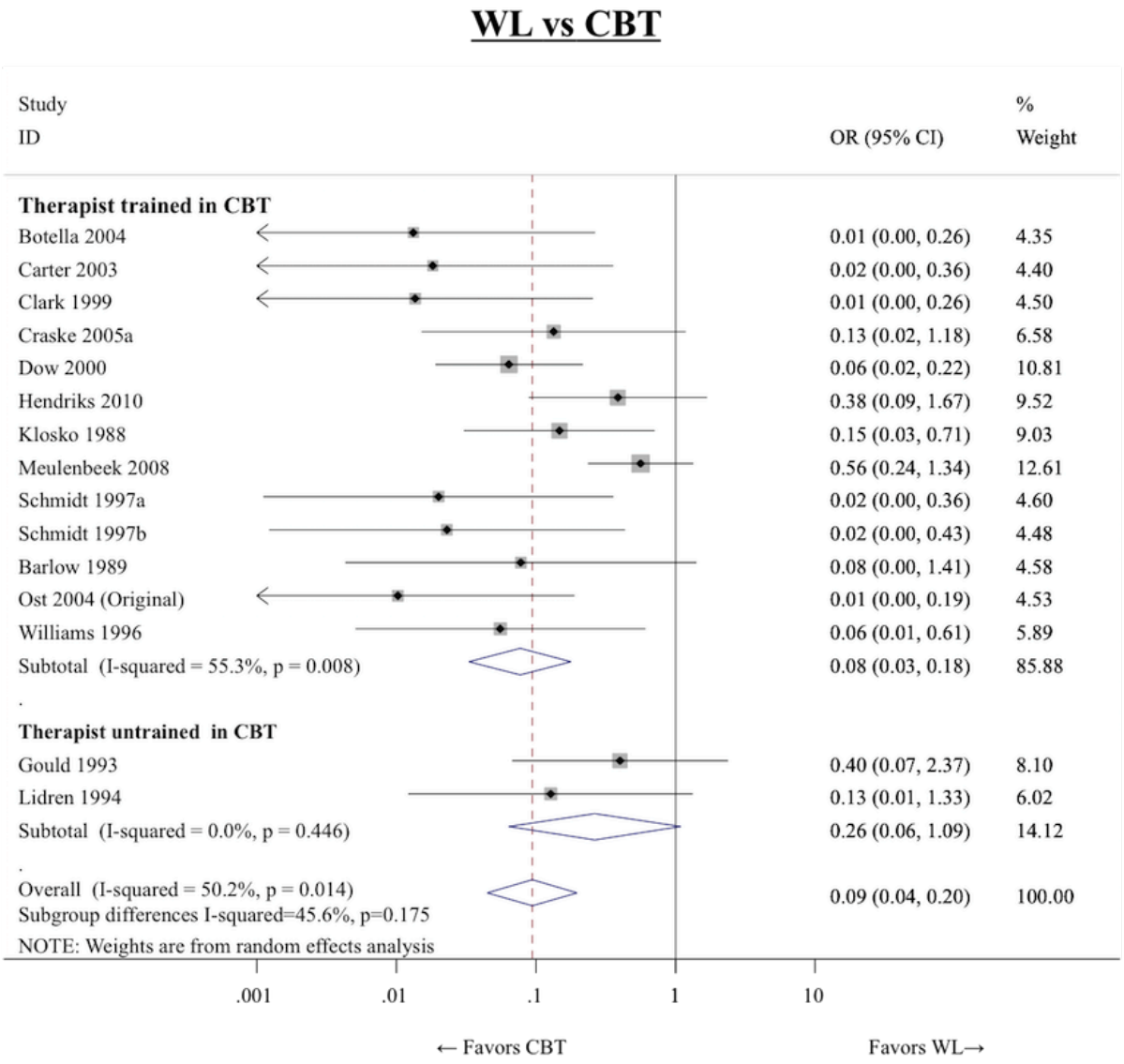


6.3 Therapist training

Following the protocol, we divided the included studies into groups depending on the assessment of therapist training in the delivered treatments. We performed a subgroup analysis for comparisons reported by enough studies (WL versus CBT and CBT versus BT).

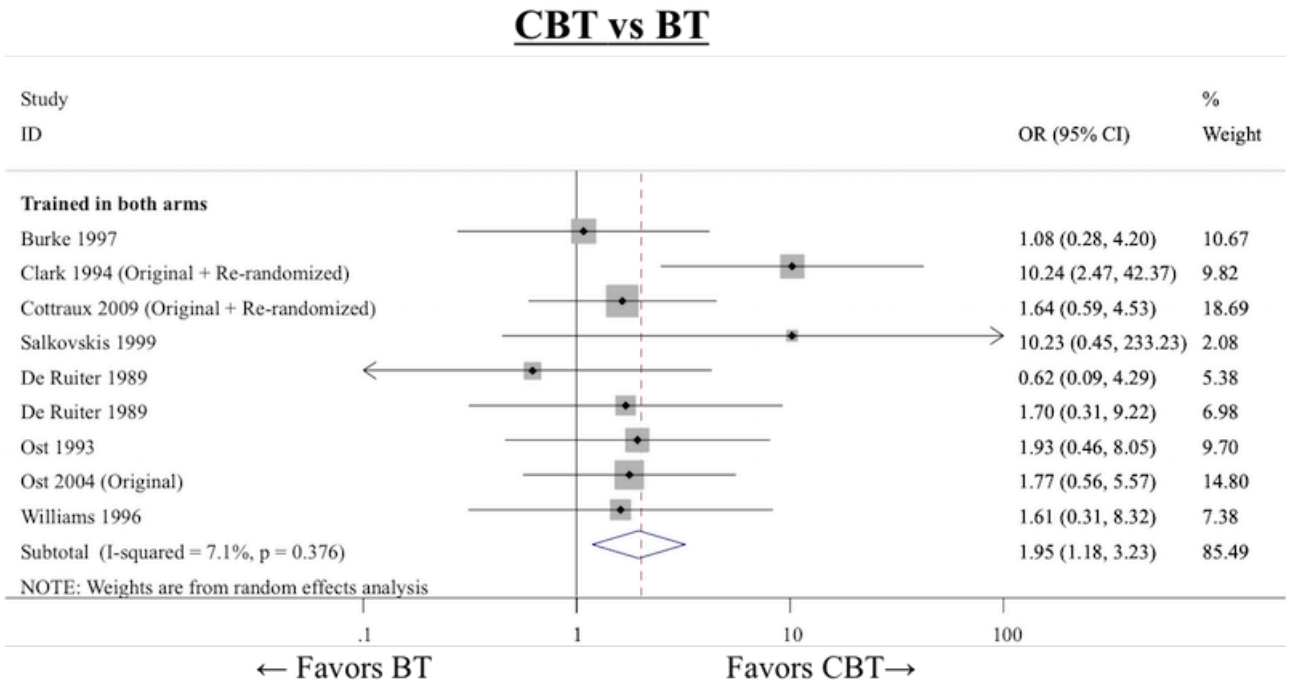
For the comparison WL versus CBT we divided the studies into two groups depending on the assessment of therapist training on CBT: therapist trained (group 1) and therapist untrained (group 2). We did not include in the analysis studies that were unclear about therapist training. The forest plots (Figure 36) showed an overlap of the confidence intervals among the groups. We thus concluded that our data showed no proof that therapist training affects the comparison between CBT and WL (test for subgroup differences: P value = 0.175; I^2 = 45.6%).

Figure 36. Subgroup analysis: therapist training, forest plots for the comparison WL vs CBT



For the comparison CBT versus BT we found three groups of studies: therapists trained in both treatments (group 1), therapists trained in BT with unclear training in CBT (group 2, one study only) and therapists with unclear training in both arms (group 3, one study only). We could only analyse the first group (Figure 37).

Figure 37. Subgroup analysis: therapist training, forest plot for the comparison CBT vs BT

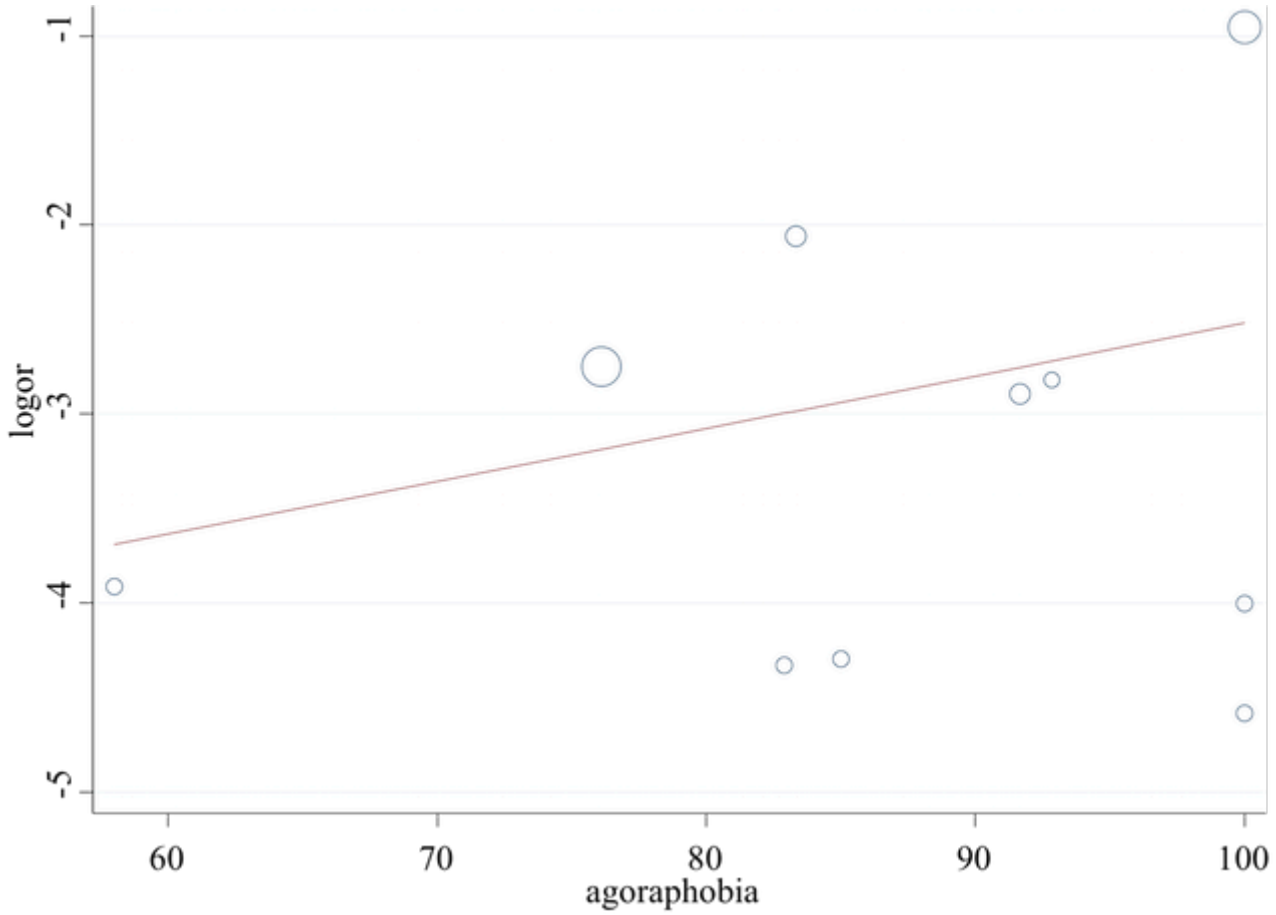


6.4 Percentage of patients with agoraphobia

We performed a subgroup analysis for comparisons reported by enough studies (WL versus CBT and CBT versus BT).

For the comparison WL versus CBT there were 10 studies reporting on agoraphobia: we found no statistically significant linear dependency between the percentage of patients with agoraphobia and the effect size for this comparison (Figure 38).

Figure 38. Subgroup analysis: percentage of agoraphobic patients, regression line for the comparison WL vs CBT



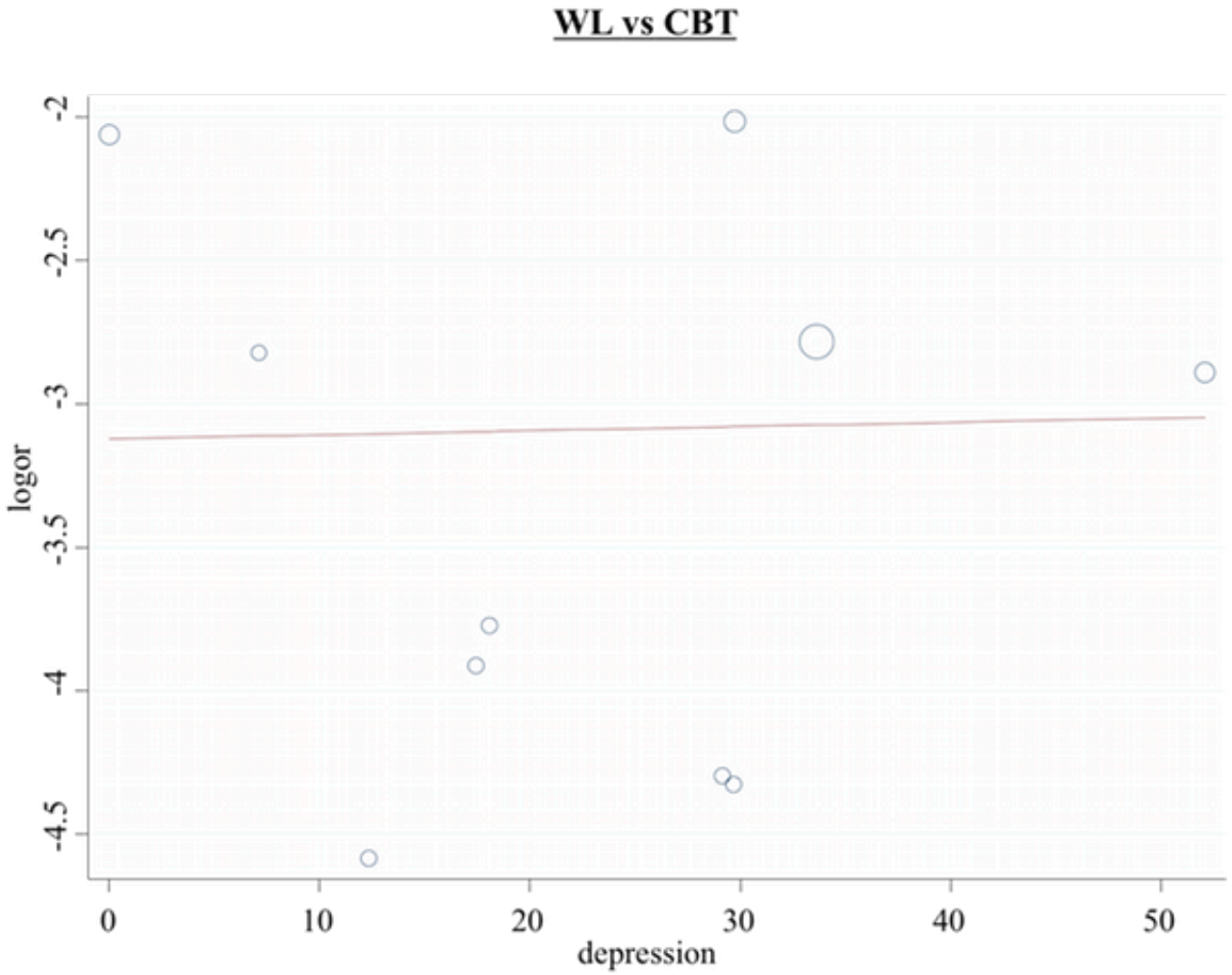
For the comparison CBT versus BT there were nine studies reporting on agoraphobia: all studies had 100% of patients with agoraphobia except one, so we could not perform an informative meta-regression for this comparison.

6.5 Percentage of patients with depression

We performed a subgroup analysis for comparisons reported by enough studies (WL versus CBT and CBT versus BT).

For the comparison WL versus CBT there were 10 studies reporting on depression: we found no statistically significant linear dependency between the percentage of patients with depression and the effect size for this comparison (Figure 39).

Figure 39. Subgroup analysis: percentage of depressed patients, regression line for the comparison WL vs CBT



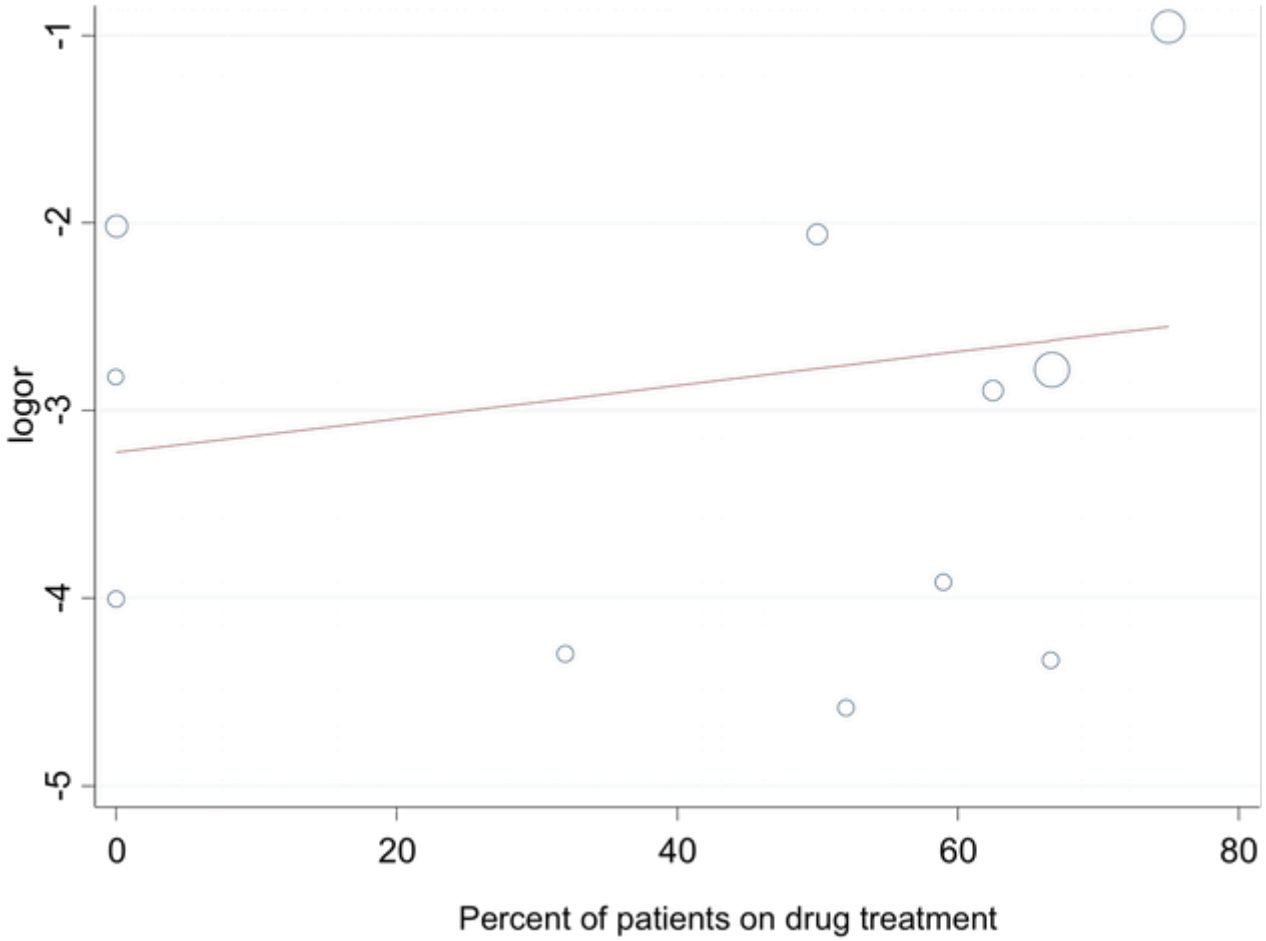
For the comparison CBT versus BT there were only four studies reporting on depression so we could not perform an informative meta-regression for this comparison.

6.6 Percentage of patients on drug treatment

We performed a subgroup analysis for comparisons reported by enough studies (WL versus CBT and CBT versus BT).

For the comparison WL versus CBT there were 11 studies reporting on drug treatment: we found no statistically significant linear dependency between the percentage of patients on drug treatment and the effect size for this comparison (Figure 40).

Figure 40. Subgroup analysis: percentage of drug-treated patients, regression line for the comparison WL vs CBT



For the comparison CBT versus BT there were only four studies reporting on drug treatment so we could not perform an informative meta-regression for this comparison.

7. Sensitivity analyses

Following the protocol, we performed sensitivity analyses for the first of the primary outcomes only (short-term remission of panic disorder with or without agoraphobia).

7.1 Analyses restricted to studies considered to be at low risk of selection and detection bias

Among the 54 studies included in quantitative analyses, only one met the criteria (Milrod 2006a). Therefore, this sensitivity analyses could not be performed.

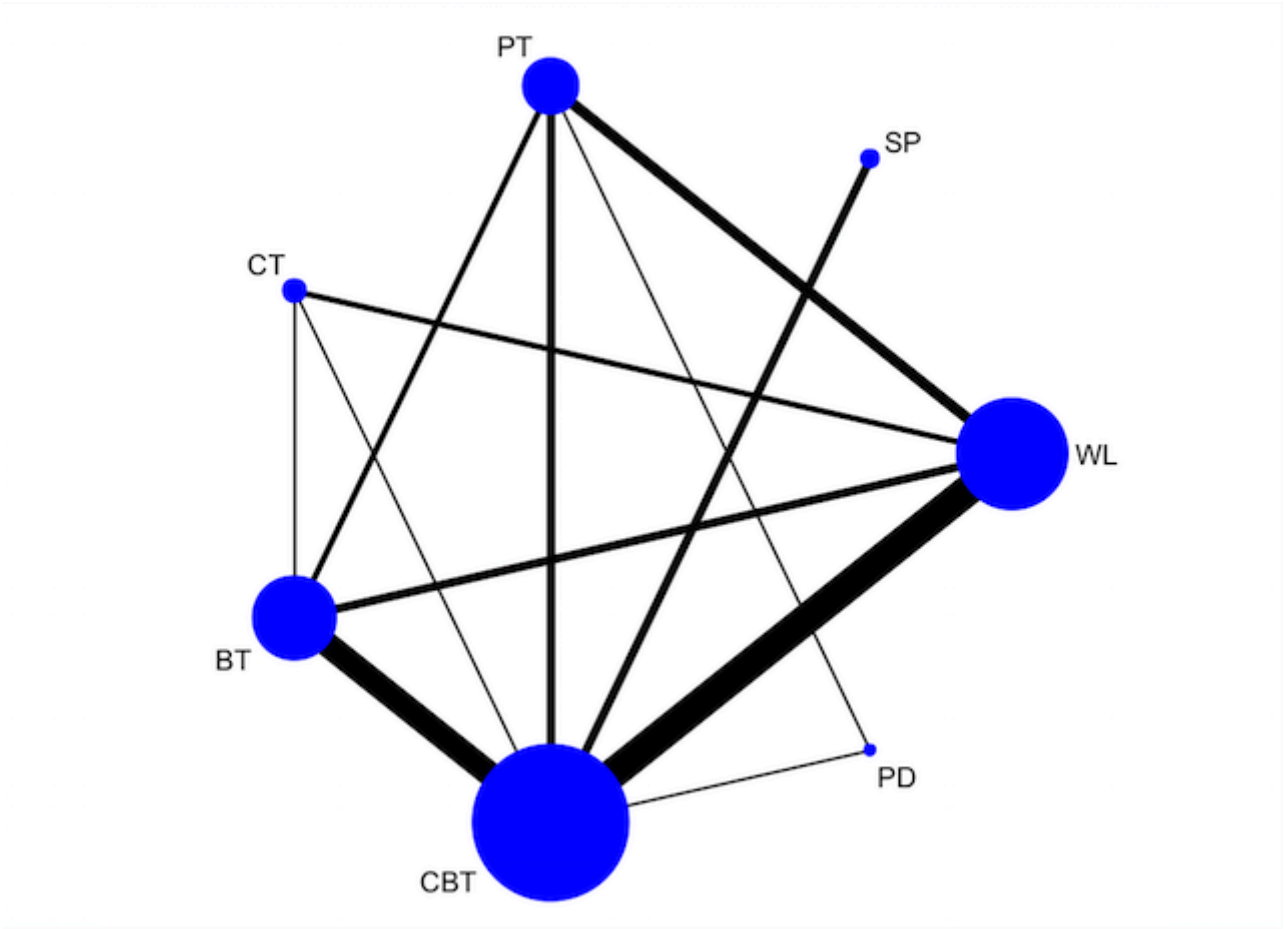
7.2 Analyses restricted to individual therapy trials

We excluded from the analysis group therapy studies and redid the analysis. We excluded 10 studies (Beck 1994; Carter 2003; Hoffart 1995; Korrelboom 2013; Lidren 1994; Meulenbeek 2008; Schmidt 1997a; Schmidt 1997b; Sharp 2004; Telch 1993).

7.2.1 Network plot

Figure 41 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, ST-remission data were available for six active and one comparison interventions. Different from the primary analyses, NT was no longer included in the network. CBT remained the most studied intervention, followed by behaviour therapy (BT), physiological therapies (PT) and cognitive therapy (CT). Wait list (WL) was the only comparator intervention available. The most studied comparison remained CBT versus WL, followed by CBT versus BT. The network remained well connected, again with the only exception of supportive psychotherapy (SP), studied only in the comparison versus CBT. Thirty studies including 1821 participants contributed data to this outcome.

Figure 41. Sensitivity analyses: network plot for short-term remission excluding from the analyses group therapy trials



7.2.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As for the primary analyses, in order to have comparable results with the NMA, we performed the pairwise meta-analyses assuming a common heterogeneity variance across all comparisons. This way, all pairwise meta-analyses were essentially analysed as random-effects (they include uncertainty due to heterogeneity), even for those comparisons only being reported by one study. The (common) heterogeneity standard deviation was estimated to be $\tau = 0.05$.

As summarised in the left part of Table 15, direct evidence was available for 12 comparisons. For four of these comparisons there was only one study available; for the remaining eight comparisons we performed a random-effects meta-analysis. As shown in the table, only one comparison was informed by 10 or more studies, that is CBT versus WL (11 studies, Figure 42); the comparison CBT versus BT was informed by nine studies, however we will present its forest plot in order to ease the comparison with primary analyses results (Figure 43). Results appear to be similar to those observed in primary analyses, with the exception of the comparison PD versus PT, which became statistically significant in this sensitivity analysis (OR 4.22, 95% CI 1.17 to 15.15), showing a superiority of PD in terms of short-term remission. However, this comparison was informed by only one study in both analyses, therefore this difference is only due to the smaller common heterogeneity standard deviation ($\tau = 0.05$ instead of 0.69) in this sensitivity analysis.

Figure 42. Sensitivity analyses: forest plot for the comparison WL vs CBT excluding group therapy trials

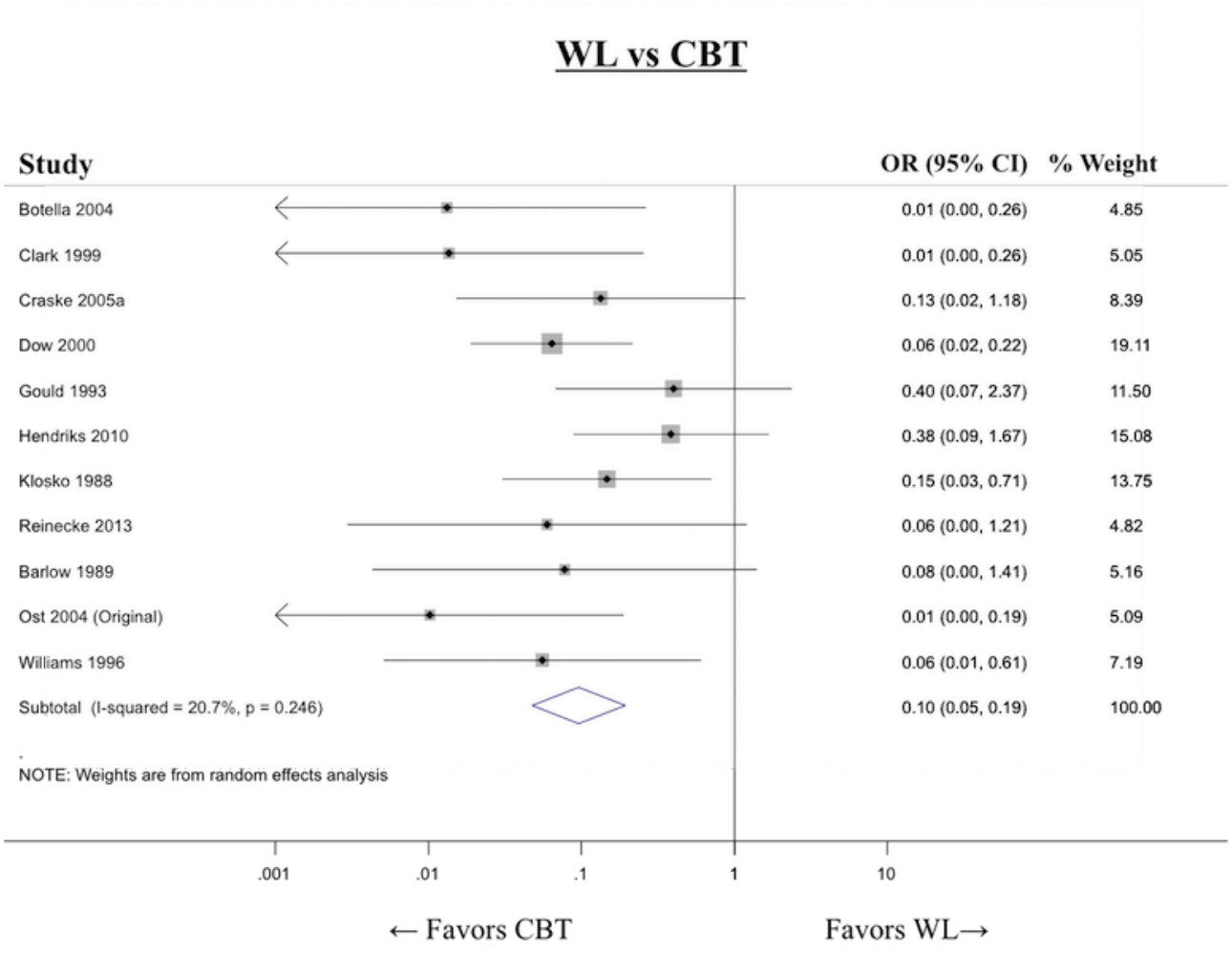
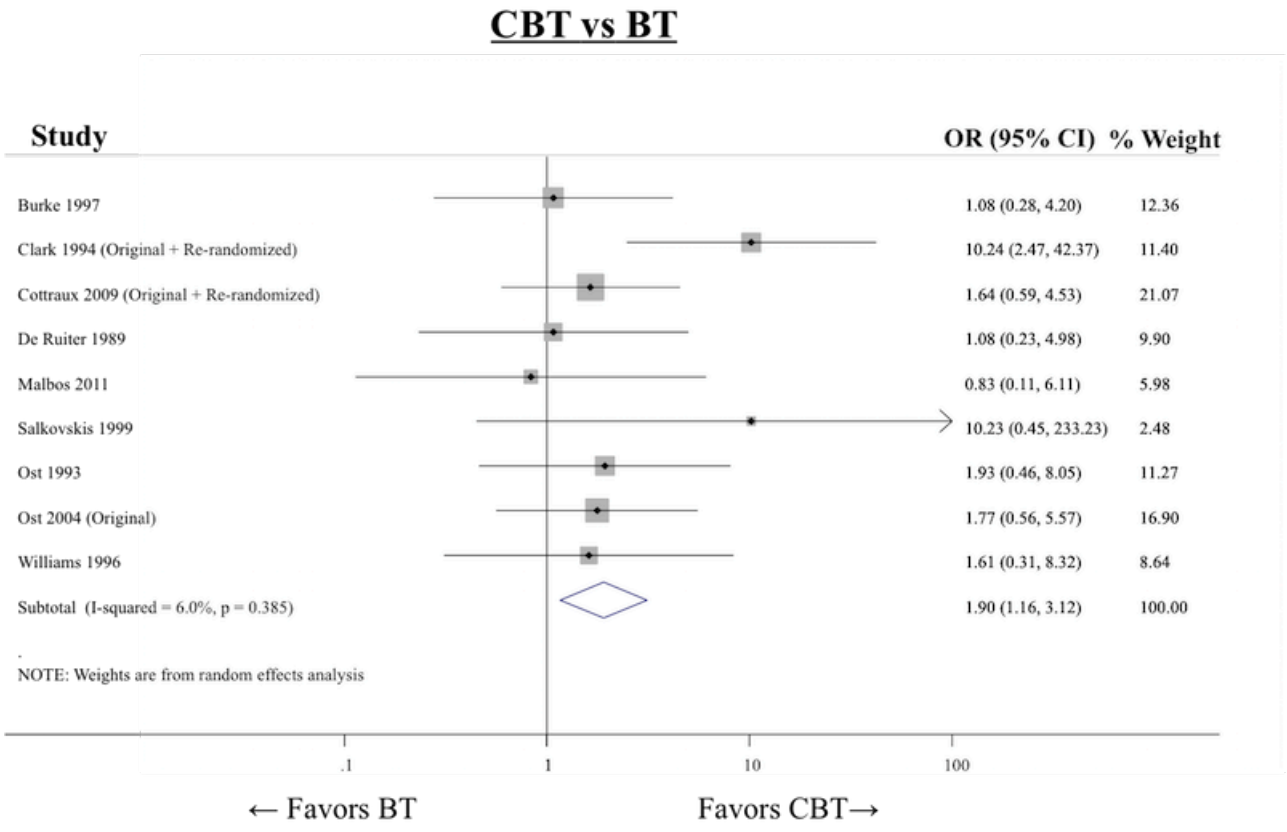


Figure 43. Sensitivity analyses: forest plot for the comparison CBT vs BT excluding group therapy trials



Heterogeneity

I^2 values and their 95% CIs, for the comparisons reported in three studies or more, are presented in Table 16. As shown in the table, we observed the highest I^2 values in the comparisons WL versus PT (I^2 = 56%). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) these values suggest that, in these two comparisons, a moderate percentage of the observed variability in the effect estimates was due to heterogeneity rather than sampling error (chance). In the comparison WL versus PT (four studies), heterogeneity was due to the study Griegel 1995, a three-arm trial (PT versus PT versus WL) in which no remission (i.e. panic-free status) was observed in the two active treatment arms whereas one case of remission was observed in the wait list. In this study, therefore, the unexpected OR was due to the low number of events across all arms.

Small study effects

We produced funnel plots for the comparisons WL versus CBT (Figure 44) and CBT versus BT (Figure 45). From the first funnel plot there was evidence of asymmetry. More specifically, small studies were missing in the lower right part of the funnel plot. This means that small studies comparing WL to CBT that (relatively) favour WL seemed to be missing: in other words, small studies showed CBT to be more efficacious. The contour-enhanced funnel plot for the comparison WL versus CBT (not presented) showed that studies were missing in the area of non-significance, thus suggesting the role of publication bias behind the SSE. We found no evidence of asymmetry in the funnel plot for the comparison CBT versus BT. Taken together, these findings are similar to those found in the primary analyses.

Figure 44. Sensitivity analyses: funnel plot for the comparison WL vs CBT excluding group therapy trials

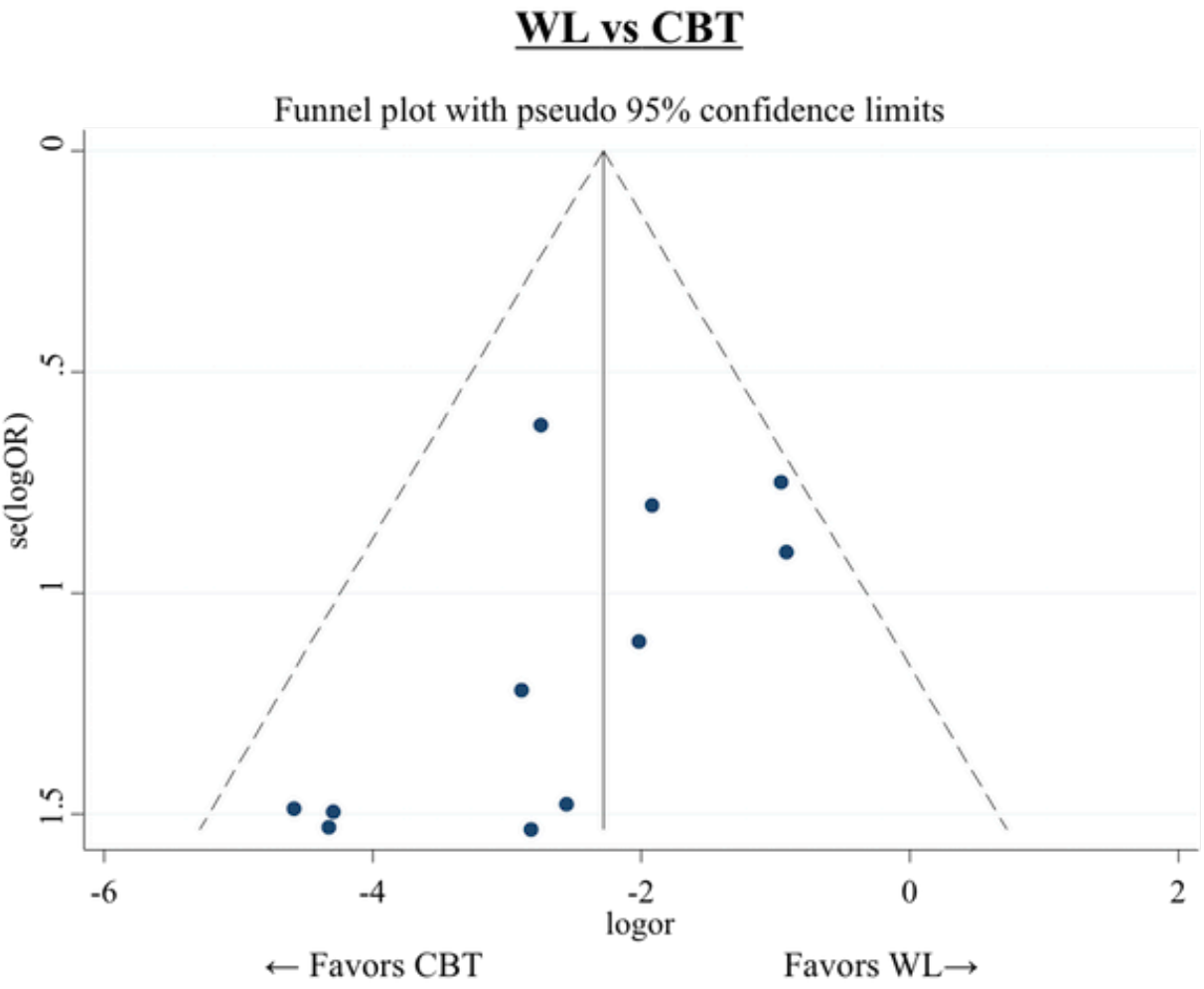
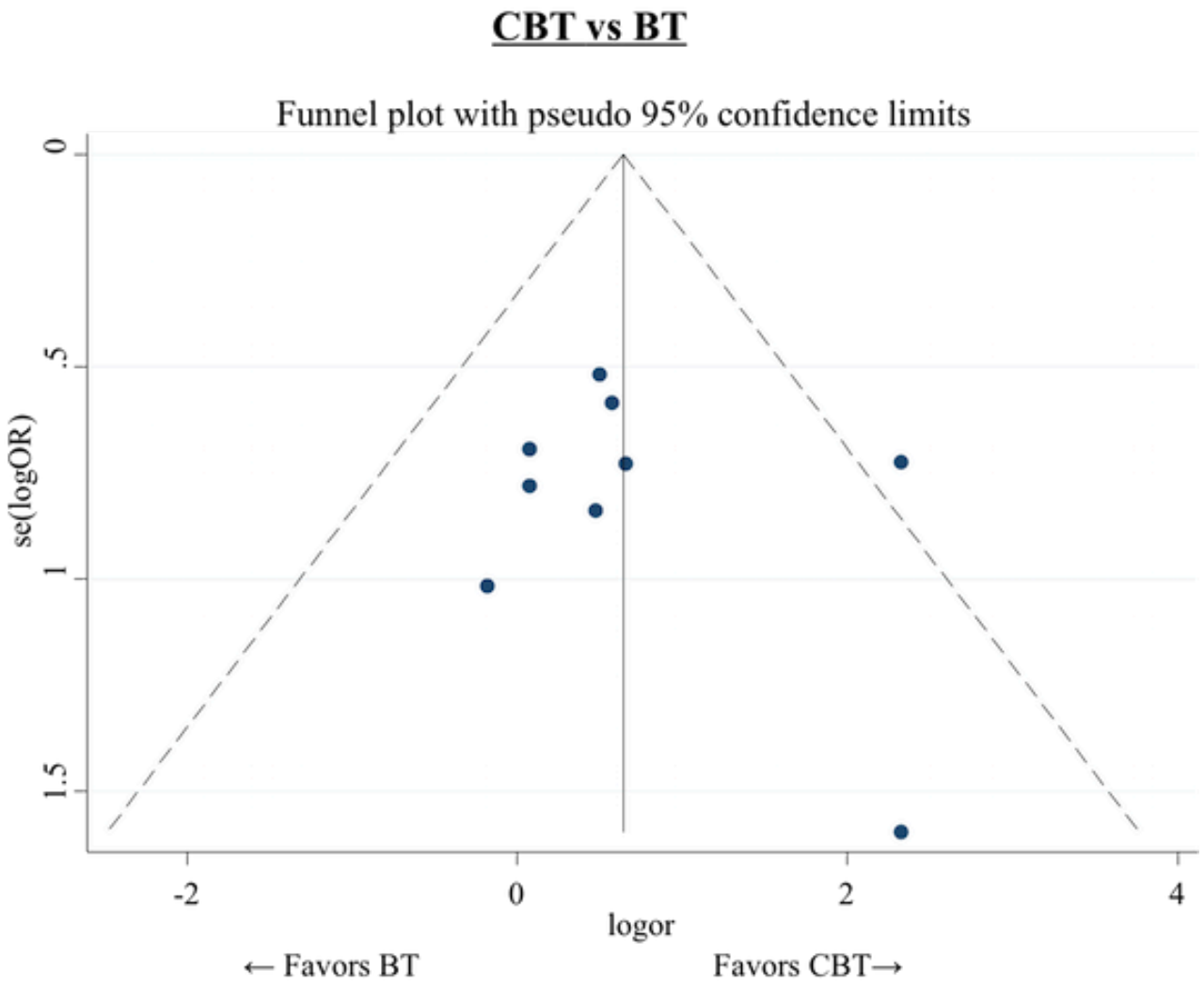


Figure 45. Sensitivity analyses: forest plot for the comparison CBT vs BT excluding group therapy trials



7.2.3 Network meta-analysis and its inconsistency

Network meta-analysis

As explained in the previous paragraph, it was evident from the funnel plots that there were small study effects (SSE) present in the network for the comparison CBT versus WL. As for the primary analyses, we found it reasonable to assume that there were SSE in all other comparisons versus WL, even though we might not have had enough studies to see this effect. The presence of SSE implies that a simple NMA may produce biased results. For this reason we performed a network meta-analysis adjusting for SSE in studies comparing all other treatments to WL, by regressing on the variance of the study. We performed the network meta-analysis adjusted for SSE in WinBUGS. Thus the results are expressed in terms of credible intervals and we use the median (instead of the mean) because the posterior distribution of the estimated odds ratios is asymmetrical.

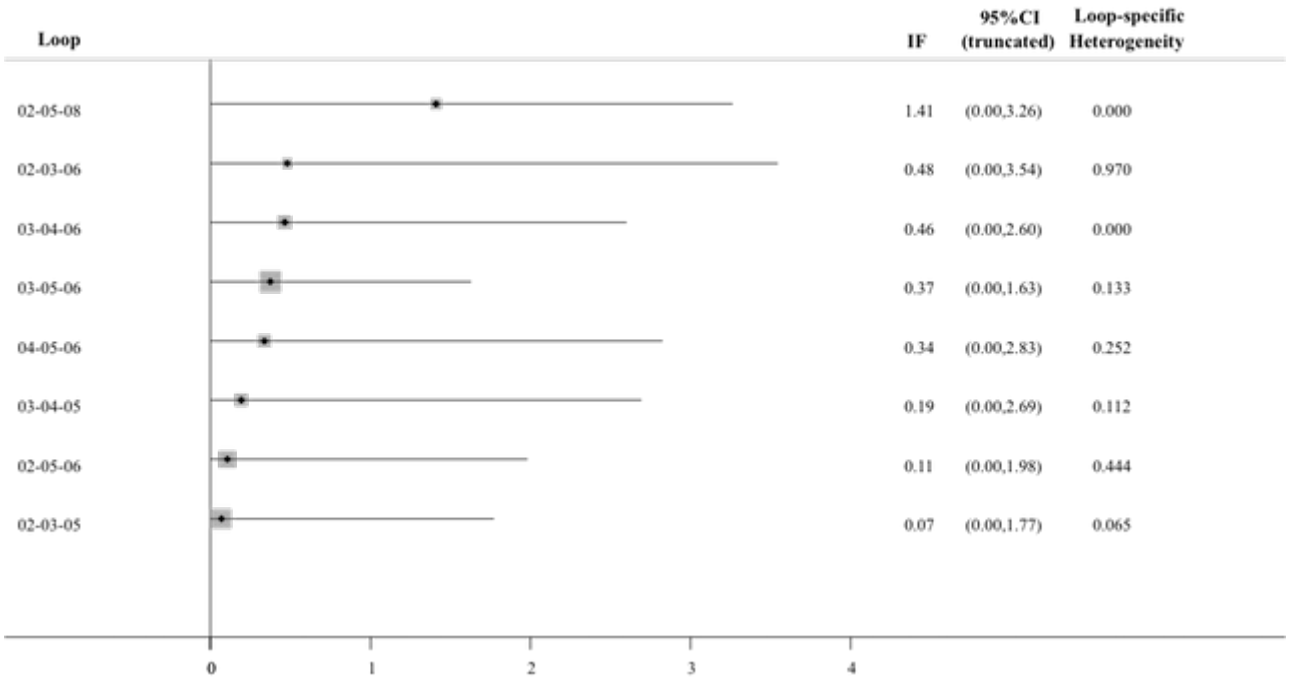
Results of the network meta-analysis (NMA), adjusted for SSE, for short-term remission are presented in the right part of Table 15. Indirect evidence could be calculated for nine comparisons for which direct evidence was unavailable. The comparison between CBT and WL remained statistically significant also within the context of NMA, showing an OR of 3.0 in favour of CBT (95% CrI 1.4 to 6.3). PT, CT and BT lost significance in the comparison versus WL. We found supportive psychotherapy (SP) to be significantly better than WL (OR 4.5, CrI 1.3 to 16.7); again, this finding must be interpreted with caution since SP is included in the network as a node with a single connection to the network. Two comparisons among active treatments, that is CBT versus BT and CBT versus PT, showed a statistically significant difference in terms of short-term remission, in both cases in favour of CBT, with an OR respectively of 1.76 (CrI 1.02 to 3.13) and 1.94 (CrI 1.02 to 3.97). Taken together, these sensitivity analyses confirmed the findings of primary analyses.

Network heterogeneity and inconsistency

The estimated value of heterogeneity lay well within the range of values usually found in Cochrane reviews, as presented by Turner 2012.

Using a loop-specific heterogeneity (where we allow the same τ for all comparisons in a loop), we found no evidence of inconsistency in the network. In Figure 46 we give all inconsistency factors for the network. As shown in the figure, the highest inconsistency factor was observed in the loop PT-CBT-PD.

Figure 46. Sensitivity analyses: inconsistency factors for the network of short-term remission excluding from the analyses group therapy trials



The global test for inconsistency showed no proof of inconsistency (Chi^2 = 3.21 with 5 degrees of freedom, P value = 0.67).

7.2.4 Ranking of treatments

The ranking of treatments with respect to short-term remission, according to the SUCRA value derived from NMA adjusted for small study effects, is presented in Table 17. Results confirmed the primary analyses, showing the highest rankings respectively for supportive psychotherapy, cognitive behaviour therapy and psychodynamic therapy. Again, results regarding supportive psychotherapy must be interpreted with caution because, as specified earlier, SP is included in the network as a node with a single connection to the network, being compared only with CBT.

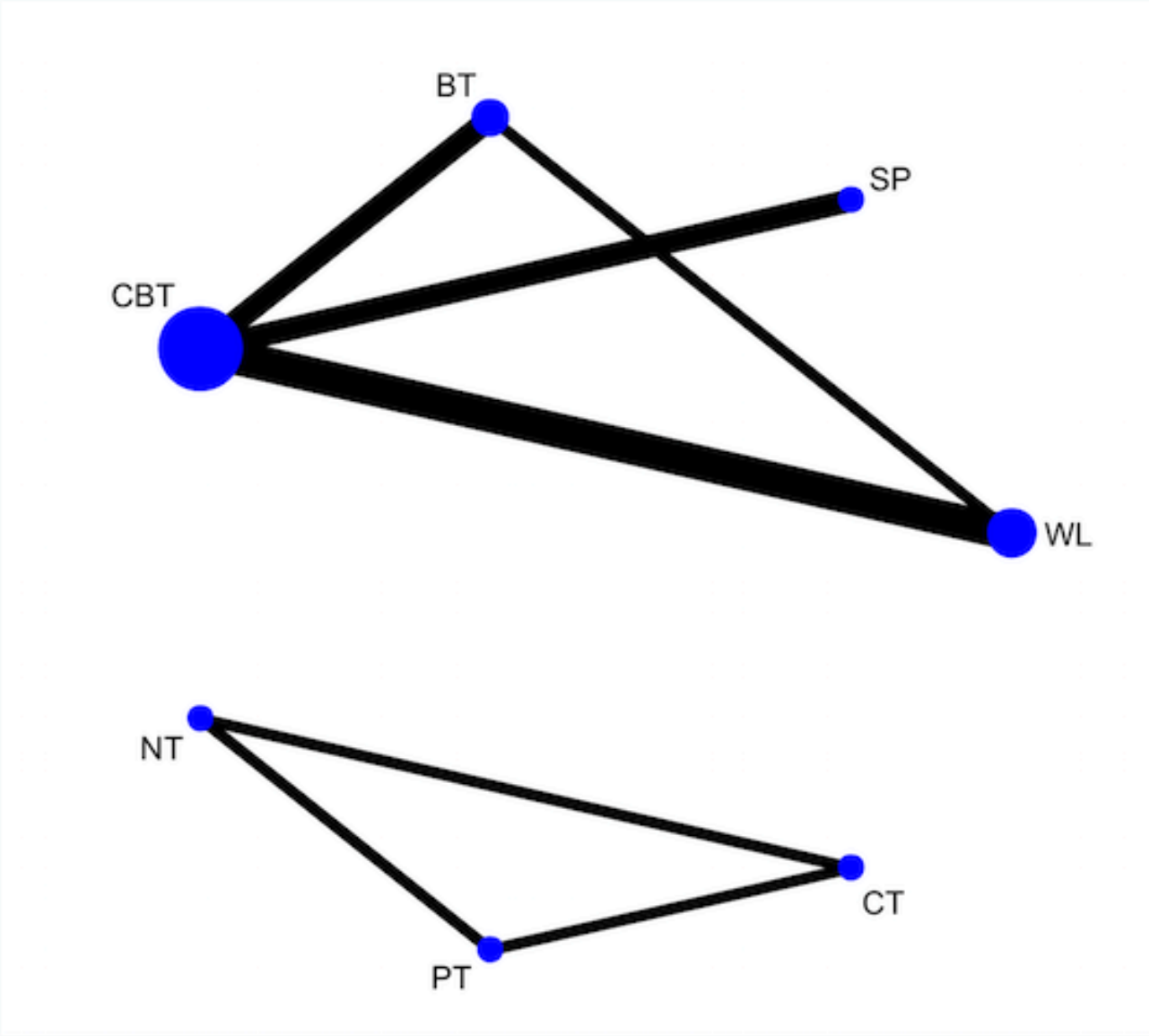
7.3 Analyses restricted to trials in which a concomitant pharmacotherapy is not allowed

We excluded from the analysis studies in which a concomitant pharmacotherapy was allowed and redid the analysis. Only nine studies remained and could be included in these analyses (Beck 1994; Carter 2003; Cottraux 2009; Craske 1995; Craske 2005a; Gloster 2010; Hoffart 1995; Reinecke 2013; Shear 1994).

7.3.1 Network plot

Figure 47 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, ST-remission data were available for five active and two comparison interventions. Different from the primary analyses, the network was no longer connected, but consisted in two separated sub-networks.

Figure 47. Sensitivity analyses: network plot for short-term remission excluding from the analyses trials in which a concomitant pharmacotherapy is allowed



7.3.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As summarised in Table 18, direct evidence was available for seven comparisons. For four of these comparisons there was only one study available; for the remaining three comparisons we performed a random-effects meta-analysis. As shown in the table, the most represented comparison was CBT versus WL (three studies). Results appear to be similar to those observed in primary analyses. The main difference was represented by the comparison WL versus CBT, which showed a larger effect (but also a larger CI) in favour of CBT compared with the primary analyses (OR 14.3, 95% CI 3.3 to 100).

Heterogeneity

I^2 values, for the comparisons reported in two studies or more, are presented in Table 18. As shown in the table, we found no heterogeneity in the small sub-samples of studies available for these sensitivity analyses.

Small study effects

Since three or fewer studies were available for each comparison, we could not explore the presence of small study effects.

7.3.3 Network meta-analysis and its inconsistency

Network meta-analysis

A network meta-analysis could not be performed because the network was disconnected.

7.4 Analyses restricted to trials requiring stabilisation of drug therapy

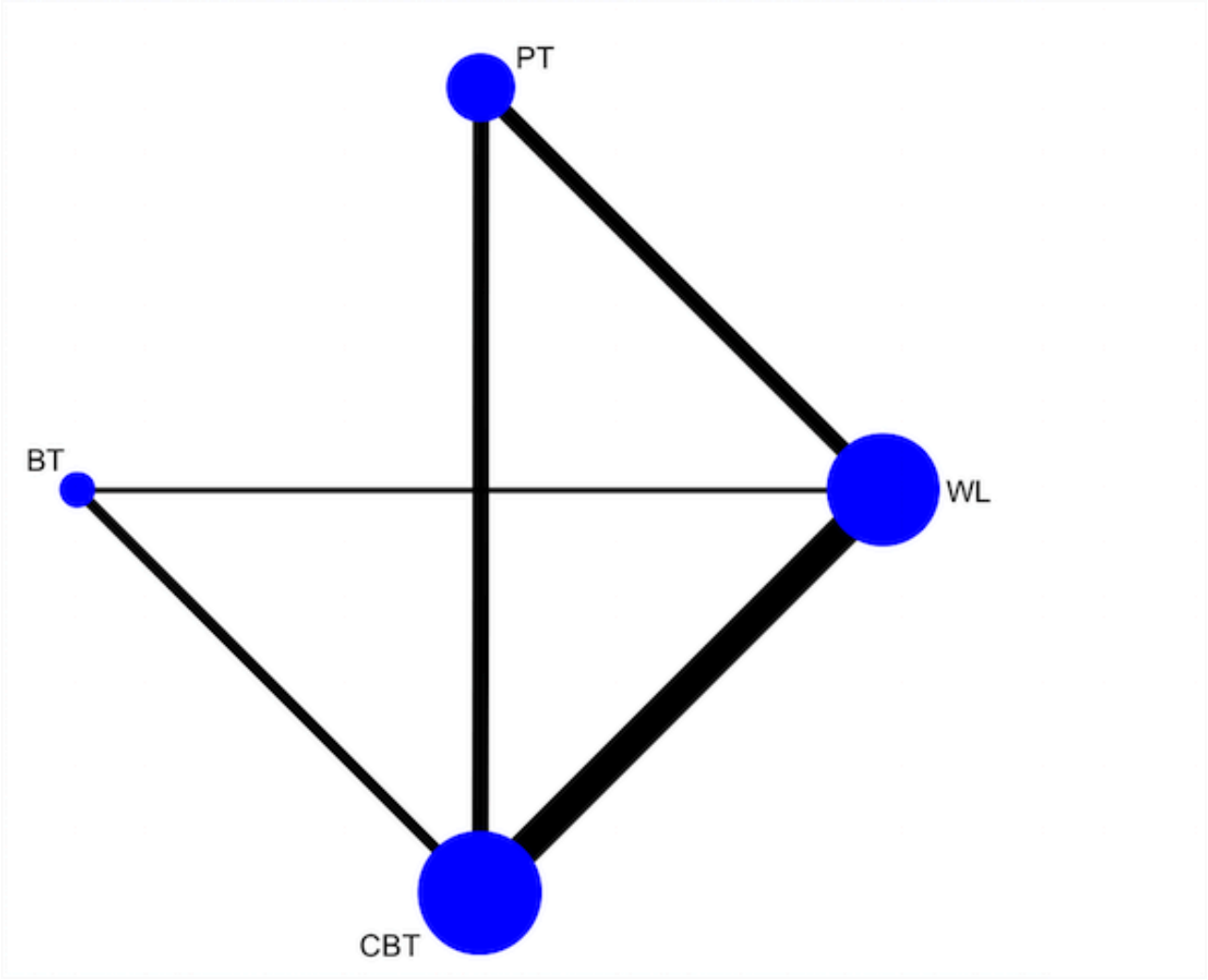
We excluded from the analysis studies in which stabilisation of drug therapy was not explicitly required and redid the analysis.

Only 11 studies remained and could be included in these analyses (Barlow 1989; Clark 1994; Craske 2005a; Dow 2000; Griegel 1995; Korrelboom 2013; Lidren 1994; Meuret 2008; Ost 1995; Ost 2004).

7.4.1 Network plot

Figure 48 shows a graphical representation of the network. Nodes and edges were weighted according to the number of studies including the respective treatments and comparisons. As shown in the figure, ST-remission data were available for three active and one comparison interventions. Although with only a few treatments were included, the network appeared to be well connected, with direct evidence missing only for the comparison BT versus PT.

Figure 48. Sensitivity analyses: network plot for short-term remission excluding from the analyses trials in which pharmacotherapy stabilisation was not required



7.4.2 Pairwise meta-analyses and their heterogeneity and small study effects

Pairwise meta-analyses

As summarised in Table 19, direct evidence was available for five comparisons. For one of these comparisons there was only one study available; for the remaining four comparisons we performed a random-effects meta-analysis. As shown in the table, the most represented comparison was CBT versus WL (six studies). Different from the primary analyses, the comparison CBT versus BT was no longer significant, although a trend in favour of CBT remained. The comparisons CBT versus WL and BT versus WL remained statistically significant, showing larger effects in favour of the active treatment with respect to primary analyses.

Heterogeneity

I^2 values, for the comparisons reported in two studies or more, are presented in Table 19. As shown in the table, we observed the highest I^2 values in the comparisons CBT versus BT (I^2 = 72%) and PT versus WL (I^2 = 69%). According to the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, section 9.5.2) these values suggest that, in these two comparisons, a substantial percentage of the observed variability in the effect estimates was due to heterogeneity rather than sampling error (chance). The low number of studies available for these comparisons does not

allow us to make inferences about the nature of the observed heterogeneity.

Small study effects

Since six or fewer studies were available for each comparison, we could not explore the presence of small study effects.

7.4.3 Network meta-analysis and its inconsistency

Results of the network meta-analysis (NMA) for short-term remission are presented in the right part of Table 19. Indirect evidence could be calculated for the only comparison for which direct evidence was unavailable. The comparisons CBT versus WL (OR 16.7, 95% CI 6.3 to 50) and BT versus WL (OR 5.3, 95% CI 1.1 to 25) remained statistically significant also within the context of NMA. Furthermore, the comparison WL versus PT became statistically significant, showing an OR of 9.09 (95% CI 2.7 to 33.3).

Compared with the primary analyses, these results tended to show larger (and statistically significant) effects of active treatments when compared with WL. Comparisons between active treatments seemed to be less affected by the exclusion of studies in which the stabilisation of drug therapy was not required.

Network heterogeneity and inconsistency

The estimated value of heterogeneity lay well within the range of values usually found in Cochrane reviews, as presented by Turner 2012.

We found no evidence of inconsistency in the network using a loop-specific approach.

The global test for inconsistency also showed no proof of inconsistency ($Chi^2 = 3.21$ with 5 degrees of freedom, P value = 0.67).

7.4.4 Ranking of treatments

The ranking of treatments with respect to short-term remission, according to the SUCRA value derived from NMA adjusted for small study effects, was CBT (SUCRA 95), PT (SUCRA 62), BT (SUCRA 42) and WL (SUCRA 0). With regard to the treatments included in this sensitivity analysis, the results were similar to those observed in the primary analyses.

7.5 Pairwise meta-analyses performed using a fixed-effect model instead of a random-effects model

For these analyses we considered the same 40 studies included in the primary analyses. As summarised in Table 20, direct evidence was available for 15 comparisons, for which we performed a fixed-effect meta-analysis and compared with the random-effects meta-analyses results. With respect to primary analyses, we observed the main difference in the comparison CBT versus WL. As we pointed out in results section 1.2, this comparison is biased in favour of CBT by the presence of small study effects; as expected, the magnitude of the effect size for this comparison was reduced using the fixed-effect model, which gives a relatively lower weight to small studies compared to the random-effects model, thus reducing their influence on meta-analysis results. A similar explanation can also account for the effect size reduction observed in the comparison BT versus WL.

DISCUSSION

Summary of main results

Overall results in terms of short-term (ST)-remission, ST-response and ST-improvement on a continuous scale were similar, suggesting a superiority of cognitive behaviour therapy (CBT), psychodynamic therapies (PD) and supportive psychotherapy (SP) over other treatments for the management of the acute phase of panic disorder.

However, results concerning SP should be interpreted cautiously because the efficacy of this treatment was explored only in the comparison with CBT and thus was not included in any closed loop of the network (in which SP appeared as a node with a single connection). As a result, the ranking of this treatment was strongly influenced by this unique available comparison, which was directly explored in three studies (Addis 2004; Craske 2005a; Shear 1994), none of which found a statistically significant difference between the two treatments in terms of ST-remission and ST-response. We suspect that this situation may have produced a spuriously high ranking of SP, because the estimation of the relative treatment effects of CBT versus SP was only informed by these three trials (and no additional indirect evidence from the rest of the network) while the comparison of SP versus all other treatments was only informed by indirect evidence (via the three CBT versus SP trials).

The two remaining treatments, that is PD and CBT, were directly compared in only one study (Beutel 2013), which suggested a superiority of CBT over PD, although the results were not statistically significant either in terms of ST-remission or in terms of ST-response (it must be noted that the lack of significance may be due to the relatively small sample size).

In terms of ST-dropouts, PD ranked higher than CBT (odds ratio (OR) 0.52, 95% confidence interval (CI) 0.15 to 1.8); a high ranking was also achieved by third-wave CBT (3W), suggesting a possible better tolerability of PD and 3W over other psychological treatments in the short term.

In the long term, CBT showed the highest ranking, followed by PD, suggesting that the effects of these two treatments may be more stable with respect to other psychological interventions. As for the short-term outcomes, only one study explored the direct comparison between PD and CBT in the long term with regard to remission, showing comparable rates of remission at six months follow-up for the two treatments (Beutel 2013); however, these results should be interpreted while taking into account the relatively high number of dropouts (almost 30% of the original sample). The superiority of CBT in the long term may be due to the administration of the so-called "relapse prevention" psychological component; however, we did not explore this specific issue and a separate review should be run in order to precisely evaluate the effects of this therapeutic component for the long-term management of the disorder. It must also be noted that in this review we did not explore the effects of drug therapy co-administration (and its adherence) on long-term remission/response, as this type of secondary analysis was limited to the first of our primary outcomes only (for the reasons explained in Subgroup analysis and investigation of heterogeneity). Taken together, the relative scarcity of long-term evidence (with respect to short-term data), the high number of dropouts at long-term assessments and the above-mentioned lack of analyses regarding

concomitant drug treatment, limit the reliability of our findings regarding long-term remission/response, which therefore should be interpreted with caution.

More generally, the results showed a statistically significant superiority of psychological therapies over the wait list condition. In particular: SP, PT, behaviour therapy (BT) and CBT were superior to wait list (WL) both in terms of ST-remission and in terms of ST-response. It must be noted, however, that the relative efficacy of psychological therapies over WL was found to be consistently affected by small study effects (SSE), as shown in the two funnel plots comparing CBT versus WL (the only available comparison versus WL for which there was a sufficient number of studies to build a funnel plot) in relation to ST-remission and ST-response. In the presence of SSE, studies with a lower standard error (i.e. larger sample size and higher number of events) tend to show smaller effects (in the case of 'positive' outcomes) than studies with a higher standard error (Sterne 2000), which implies that the latter may lead to an overestimation of the effect (in this case, an overestimation of ST-remission and ST-response of CBT when compared to WL). Since the presence of SSE is difficult to detect when only a few studies are available for a comparison, we found it reasonable to assume that what we clearly observed in the comparison CBT versus WL could probably be extended to other comparisons versus WL. In other words, we assumed that there were SSE in all other comparisons versus WL, even though we might not have had enough studies to see this effect. Since the presence of SSE implies that a simple network meta-analysis (NMA) may produce biased results, we decided to repeat the NMA analyses adjusting for small study effects in studies comparing all other treatments to WL, by regressing on the variance of the study. As suspected, in adjusted analyses, many of the comparisons mentioned above lost statistical significance (only CBT and SP remained superior to WL in terms of ST-remission; no treatment remained superior in terms of ST-response).

The results showed a statistically significant difference between two active treatments only for the comparisons CBT versus BT and CBT versus PT, where CBT appeared to be superior in terms of ST-remission. However, the confidence interval was large and its lower end was very close to 1 (i.e. no difference) for both comparisons, which limits the relevance of these findings in clinical terms. A similar trend in favour of CBT over BT and PT was found in terms of ST-response, ST-improvement and LT-remission/response, although this was not statistically significant.

As planned in the protocol, we produced three 'Summary of findings' tables presenting the NMA results for the comparison between the psychological therapy that ranked first versus, respectively: no treatment (NT), supportive psychotherapy (SP) and the psychological therapy that ranked second. The ranking we referred to was the one related to the first of our primary outcomes (short-term remission), presented in Table 3. According to these results, the treatment that ranked first was supportive psychotherapy; however, for the reasons explained above, results concerning SP were not reliable, which left CBT as the treatment with the highest ranking. Therefore, in the 'Summary of findings' tables we present the NMA results for the following comparisons: CBT versus NT (Summary of findings for the main comparison), CBT versus SP (Summary of findings 2) and CBT versus PD (Summary of findings 3). We also produced an extra summary of findings table, not planned in the protocol, in order to summarize the overall results of network meta-analyses in terms of treatment hierarchy, together with the corresponding assessment of quality of the evidence (Summary of findings 4).

Overall completeness and applicability of evidence

Sufficient evidence supports our findings in relation to the comparison between CBT and waiting list: despite the evidence of small study effects affecting results for ST-remission and ST-response in favour of CBT, the stability of results when adjusting NMA for SSE suggests a good reliability of this finding. We found evidence in relation to the comparison BT versus CBT, supporting the superiority of the latter. Some evidence exists in support of the possible viability of psychodynamic and supportive psychological therapies as valid alternatives to CBT for the treatment of panic disorder with or without agoraphobia; scarce evidence exists in support of other psychological therapies as valid alternatives to CBT.

Quality of the evidence

For the three comparisons presented in the SoF tables (CBT vs NT, CBT vs SP, CBT vs PD), the quality of evidence was rated as low to very low for ST-remission, very low for ST-response, very low for ST-dropouts, low for LT-remission/response and low to very low for ST-improvement as measured on a continuous scale (Summary of findings for the main comparison; Summary of findings 2; Summary of findings 3).

We found low quality evidence, for each of the included outcomes, in support of NMA analyses regarding treatment hierarchy (Summary of findings 4).

The main factor affecting the quality of evidence was the presence of an unclear or high risk of bias, for many included studies, in more than one important domain. As shown in Figure 3 and Figure 4, the majority of included studies were at unclear risk of bias with regard to the randomisation process; we found almost half of the included studies to be at high risk of attrition bias (both in the short and in the long term) and detection bias (with regard to ST-remission); we also found selective outcome reporting to be present. Finally, we found almost half of the included studies to be at high risk of researcher allegiance bias: however, we found no clear evidence of this bias affecting our findings (extra analyses, not reported).

We strongly suspected the presence of publication bias based on evidence of SSE and analysis of contour-enhanced funnel plots for both ST-remission and ST-response, for the comparisons between psychological therapies and WL. As a post-hoc analysis we also performed the Harbord and Peters tests, which showed no proof of small study effects. However, given that the funnel plots provided strong evidence for publication bias, we employed a hierarchical network meta-regression model to adjust the NMA results (for this reason we decided not to downgrade the quality of evidence because of publication bias).

Another factor affecting the quality of evidence was the lack of precision of results for many comparisons, which in many cases did not show statistically significant differences: this was probably due to the lack of enough (and adequately powered) studies exploring such comparisons.

Despite the above mentioned limitations of available evidence, we found substantial heterogeneity to be present in very few pairwise

comparisons, limited to one of the secondary outcomes, that is ST-improvement as measured on a continuous scale. For this same outcome, the network analyses revealed the presence of relevant inconsistency in one loop (PT-CBT-PD). This was the only clear case of inconsistency identified in our analyses, although it must be said that the networks were generally underpowered to detect any important disagreement between direct and indirect evidence.

Finally, with regard to the directness of evidence, we found the networks to be moderately connected, which means that direct evidence was available for about half of the possible comparisons among included psychological therapies/control conditions. Only in the case of ST-dropouts results were informed by a higher proportion of indirect evidence; on the other hand, however, this outcome was informed by the highest number of studies (47 RCTs, 2535 participants).

Potential biases in the review process

Although this is the largest and most comprehensive systematic review and network meta-analysis of psychological therapies for panic disorder, there is reason to suspect that the available literature may be affected by publication bias. Even when a study was published, lack of unified outcome measures across studies, especially in the older trials, left suspicion of outcome reporting bias. Many of the included studies were not ideal in terms of risk of bias as individual studies. Our comprehensive literature search, unified and reliable data extraction and study assessment, and methodological rigour in the analyses have been able to guard against some but not all of these limitations.

Agreements and disagreements with other studies or reviews

To our knowledge, this is the most extensive and methodologically rigorous systematic review ever run on psychological therapies for the treatment of panic disorder. Furthermore, this is the first time in which such therapies have been compared with each other through a network meta-analysis.

Our results confirm the findings of a previous pairwise meta-analysis (Mitte 2005), in which CBT was found to be superior to waiting list in the reduction of anxiety and depressive symptoms and in the improvement of quality of life. However, in Mitte 2005 no difference was found between CBT and BT in terms of anxiety reduction. The results of a more recent meta-analysis suggested that the combination of exposure, relaxation/breathing techniques and cognitive therapy may represent the most effective treatment for panic disorder, with smaller effect sizes for any of these components if administered alone (Sánchez-Meca 2010). This finding is more in line with our results, where CBT appeared to perform better than CT, BT and PT.

Different from Sánchez-Meca 2010 and Mitte 2005, where publication bias was discarded as a possible threat to the validity of results, we found clear evidence of small study effects (which may be due, at least partly, to publication bias) affecting the comparison of CBT versus waiting list.

In Mitte 2005, the author hypothesised that the difference in effect sizes found between the comparisons (C)BT versus wait list and (C)BT versus placebo suggested a relatively large common factor, explaining more than half of the efficacy. In another meta-analysis CBT was compared with placebo (both psychological and pill placebo) for various anxiety disorders, among which the smallest effect size was observed for panic disorder (Hedges' g = 0.35, 95% CI 0.04 to 0.65), suggesting again that non-specific factors may play an important role in the treatment of this disorder (Hofmann 2008). Although psychological placebo and supportive psychotherapy are deeply different concepts, it is often difficult to draw a clear line between them in psychotherapy trials. The evidence we found in favour of supportive psychotherapy could therefore be considered in line with the findings of above mentioned reviews, suggesting that non-specific factors may actually play an important role in the treatment of panic disorder.

AUTHORS' CONCLUSIONS

Implications for practice

There was no high-quality, unequivocal evidence to support one psychological therapy over the others for the treatment of panic disorder with or without agoraphobia in adults. The ranking of treatments according to the SUCRA value, derived from NMAs, showed that cognitive behaviour therapy (CBT), the most extensively studied among the included psychological therapies, was often superior to other therapies, although the effect size was small and the level of precision was often insufficient or clinically irrelevant.

In the only two available studies exploring psychodynamic psychotherapies, this treatment showed promising results, although further research is needed in order to better explore the relative efficacy of psychodynamic therapies (PD) with respect to CBT. Unexpectedly, we found some evidence in support of the possible viability of non-specific supportive psychotherapy for the treatment of panic disorder; however, the results concerning supportive psychotherapy (SP) should be interpreted cautiously because of the sparsity of evidence regarding this treatment and, as was the case for PD, further research is needed to explore this issue. Behaviour therapy did not appear to be a valid alternative to CBT as a first-line treatment for patients with panic disorder with or without agoraphobia.

Implications for research

An important finding of this review regards the quality level of trials. It is desirable that future trials present more detailed descriptions of the randomisation process, ensure the blinding of outcome assessors and provide an a priori specification of primary and secondary outcomes.

There seems to be no further need to explore the comparison between CBT and wait list (WL), nor between CBT and behaviour therapy (BT); rather, there is a need for studies exploring the comparison between active treatments, with particular regard to CBT, PD and supportive psychotherapy, possibly in the context of multi-arm trials, with large sample sizes, including long-term assessments.

REFERENCES

References to studies included in this review

Addis 2004 *{published data only}*

Addis ME, Hatgis C, Cardemil E, Jacob K, Krasnow AD, Mansfield A. Effectiveness of cognitive-behavioral treatment for panic disorder versus treatment as usual in a managed care setting: 2-year follow-up. *Journal of Consulting and Clinical Psychology* 2006;**74**(2):377-85.

* Addis ME, Hatgis C, Krasnow AD, Jacob K, Bourne L, Mansfield A. Effectiveness of cognitive-behavioral treatment for panic disorder versus treatment as usual in a managed care setting. *Journal of Consulting and Clinical Psychology* 2004;**72**(4):625-35.

Al Kubaisy 1992 *{published and unpublished data}*

* Al Kubaisy T, Marks I, Logsdail S, Marks MP, Lovell K, Sungur M, et al. Role of exposure homework in phobia reduction: a controlled study. *Behavior Therapy* 1992;**23**:599-621.

Park JM, Mataix-Cols D, Marks IM, Ngamthipwatthana T, Marks M, Araya R, et al. Two-year follow-up after a randomised controlled trial of self- and clinician-accompanied exposure for phobia/panic disorders. *British Journal of Psychiatry* 2001;**178**:543-8.

Arch 2012 *{published data only}*

* Arch JJ, Eifert GH, Davies C, Plumb VJC, Rose RD, Craske MG. Randomized clinical trial of cognitive behavioral therapy (CBT) versus acceptance and commitment therapy (ACT) for mixed anxiety disorders. *Journal of Consulting and Clinical Psychology* 2012;**80**(5):750-65.

Arch JJ, Wolitzky-Taylor KB, Eifert GH, Craske MG. Longitudinal treatment mediation of traditional cognitive behavioral therapy and acceptance and commitment therapy for anxiety disorders. *Behaviour Research and Therapy* 2012;**50**:469-78.

Arntz 2002 *{published data only (unpublished sought but not used)}*

Arntz A. Cognitive therapy versus interoceptive exposure as treatment of panic disorder without agoraphobia. *Behaviour Research and Therapy* 2002;**40**(3):325-41.

Barlow 1989 *{published data only (unpublished sought but not used)}*

Barlow DH, Brown TA, Craske MG, Rapee RM, Antony MM. Treatment of panic disorder: follow-up and mechanisms of action. 25th Annual Meeting of the Association for the Advancement of Behavior Therapy. New York, 1991.

* Barlow DH, Craske MG, Cerny JA, Klosko JS. Behavioral treatment of panic disorder. *Behavior Therapy* 1989;**20**(2):261-82.

Brown TA, Antony MM, Barlow DH. Diagnostic comorbidity in panic disorder: effect on treatment outcome and course of comorbid diagnoses following treatment. *Journal of Consulting and Clinical Psychology* 1995;**63**(3):408-18.

Brown TA, Barlow DH. Long-term outcome in cognitive-behavioral treatment of panic disorder: clinical predictors and alternative strategies for assessment. *Journal of Consulting and Clinical Psychology* 1995;**63**(5):754-65.

Craske MG, Brown TA, Barlow DH. Behavioral treatment of panic disorder: a two-year follow-up. *Behavior Therapy* 1991;**22**(3):289-304.

Margraf J, Barlow DH, Clark DM, Telch MJ. Psychological treatment of panic: work in progress on outcome, active ingredients, and follow-up. *Behavior Research and Therapy* 1993;**31**(1):1-8.

Beck 1987 *{published data only (unpublished sought but not used)}*

* Beck AT. Cognitive approaches to panic disorder: theory and therapy. In: Rachman S, Maseck J editor(s). Panic: Psychological Perspectives. Hillsdale, NJ: L. Erlbaum Associates, 1988:91-109.

Sokol-Kessler L, Beck AT. Cognitive treatment of panic disorders. 140th Annual Meeting of the American Psychiatric Association. Chicago, May 1987.

Beck 1992 *{published data only (unpublished sought but not used)}*

Beck AT, Sokol L, Clark DA, Berchick R, Wright F. A crossover study of focused cognitive therapy for panic disorder. *American Journal of Psychiatry* 1992;**149**(6):778-83.

Beck 1994 *{published data only}*

* Beck JG, Stanley MA, Baldwin LE, Deagle EA, Averill PM. Comparison of cognitive therapy and relaxation training for panic disorder. *Journal of Consulting and Clinical Psychology* 1994;**62**(4):818-26.

Stanley MA, Beck JG, Averill PM, Baldwin LE, Deagle EA, Stadler JG. Patterns of change during cognitive behavioral treatment for panic disorder. *Journal of Nervous and Mental Disease* 1996;**184**(9):567-72.

Beutel 2013 *{published and unpublished data}*

Beutel M. Short-term psychotherapy for patients with panic disorder with or without agoraphobia - a randomised comparative study between the panic-focused psychodynamic and cognitive behavioural psychotherapy [Kurzzeittherapie der Panikstörung mit oder ohne Agoraphobie mittels panikfokussierender Psychodynamischer Psychotherapie (PfPP) und kognitiver Verhaltenstherapie (KVT) - eine randomisierte und kontrollierte Therapievergleichsstudie]. https://drks-neu.uniklinik-freiburg.de/ 2009 (accessed 16 March 2015):S00000245. [DRKS00000245]

* Beutel ME, Scheurich V, Knebel A, Michal M, Wiltink J, Graf-Morgenstern M, et al. Implementing panic-focused psychodynamic psychotherapy into clinical practice. *Canadian Journal of Psychiatry* 2013;**58**(6):326-34.

Botella 2004 *{published and unpublished data}*

Botella C, Garcia-Palacios A, Villa H, Banos RM, Quero S, Alcaniz M, et al. Virtual reality exposure in the treatment of

panic disorder and agoraphobia: a controlled study. *Clinical Psychology and Psychotherapy* 2007;**14**(3):164-75.

* Botella C, Villa H, Garcia Palacios A, Quero S, Banos RM, Alcaniz M. The use of VR in the treatment of panic disorders and agoraphobia. *Studies in Health Technology & Informatics* 2004;**99**:73-90.

Brown 1997 *{published data only (unpublished sought but not used)}*

Brown GK, Beck AT, Newman CF, Beck JS, Tran GQ. A comparison of focused and standard cognitive therapy for panic disorder. *Journal of Anxiety Disorders* 1997;**11**(3):329-45.

Burke 1997 *{published data only}*

Burke M, Drummond LM, Johnston DW. Treatment choice for agoraphobic women: exposure or cognitive-behaviour therapy?. *British Journal of Clinical Psychology* 1997;**36**(Pt 3):409-20.

Carter 2003 *{published data only}*

Carter MM, Sbrocco T, Gore KL, Marin NW, Lewis EL. Cognitive-behavioral group therapy versus a wait-list control in the treatment of African American women with panic disorder. *Cognitive Therapy and Research* 2003;**27**(5):505-18.

Clark 1994 *{published data only}*

Clark DM, Salkovskis PM, Hackmann A, Middleton H, Anastasiades P, Gelder M. A comparison of cognitive therapy, applied relaxation and imipramine in the treatment of panic disorder. *British Journal of Psychiatry* 1994;**164**(6):759-69.

Clark 1999 *{published data only}*

Clark DM, Salkovskis PM, Hackmann A, Wells A, Ludgate J, Gelder M. Brief cognitive therapy for panic disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology* 1999;**67**(4):583-9.

Cottraux 2009 *{published data only (unpublished sought but not used)}*

Cottraux J. A comparative controlled study of virtual reality therapy and cognitive behavior therapy in panic disorder with agoraphobia. https://clinicaltrials.gov/ct2/show/NCT00129610 2005 (accessed 16 March 2015).

Lambrey S, Jouvent R, Allilaire JF, Pelissolo A. Virtual reality therapies in the treatment of phobic disorders [French]. *Annales Medico-Psychologiques* 2010;**168**(1):44-6.

* Pelissolo A, Zaoui M, Aguayo G, Yao SN, Roche S, Ecochard R, et al. Virtual reality exposure therapy versus cognitive behavior therapy for panic disorder with agoraphobia: a randomized comparison study. *Journal of CyberTherapy and Rehabilitation* 2012;**5**(1):35-43.

Pull C, Pelissolo A, Zaoui M, et al. A randomized controlled study of virtual reality exposure therapy and cognitive-behavior therapy in panic disorder with agoraphobia. *Cyberpsychology & Behavior (Abstracts from CyberTherapy 14 Designing the Future of Healthcare, June 21–23, 2009, Lago Maggiore, Verbania, Italy)* 2009;**12**(5):647.

Craske 1995 *{published data only}*

Craske MG, Maidenberg E, Bystritsky A. Brief cognitive-behavioral versus nondirective therapy for panic disorder. *Journal of Behavior Therapy and Experimental Psychiatry* 1995;**26**(2):113-20.

Craske 2005a *{published data only}*

* Craske MG, Lang AJ, Aikins D, Mystkowski JL. Cognitive behavioral therapy for nocturnal panic. *Behavior Therapy* 2005;**36**(1):43-54.

Tsao JC, Mystkowski JL, Zucker BG, Craske MG. Impact of cognitive-behavioral therapy for panic disorder on comorbidity: a controlled investigation. *Behaviour Research and Therapy* 2005;**43**(7):959-70.

Creager Berger 2001 *{published data only}*

* Creager Berger B. The treatment of panic disorder: a comparative study between breathing retraining and cognitive behavioral therapy. Dissertation Abstracts International 2001; Vol. 61, issue 8b.

Creager Berger B, Gevirtz R. The treatment of panic disorder: a comparison between breathing retraining and cognitive behavioral therapy. *Biological Psychology* 2001;**56**(1):78-9.

De Ruiter 1989 *{published data only}*

Rijken H, Kraaimaat F, de Ruiter C, Garssen B. A follow-up study on short-term treatment of agoraphobia. *Behaviour Research and Therapy* 1992;**30**(1):63-6.

* de Ruiter C, Ryken H, Garssen B, Kraaimaat F. Breathing retraining, exposure and a combination of both, in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy* 1989;**27**(6):647-55.

Dow 2000 *{published data only (unpublished sought but not used)}*

Casey LM, Newcombe PA, Oei TP. Cognitive mediation of panic severity: the role of catastrophic misinterpretation of bodily sensations and panic self-efficacy. *Cognitive Therapy and Research* 2005;**29**(2):187-200.

Dow MG, Kenardy JA, Johnston DW, Newman MG, Taylor CB, Thomson A. Prognostic indices with brief and standard CBT for panic disorder: I. Predictors of outcome. *Psychological Medicine* 2007;**37**(10):1493-502.

Kenardy J, Robinson S, Dob R. Cognitive behaviour therapy for panic disorder: long-term follow up. *Cognitive Behaviour Therapy* 2005;**34**(2):75-8.

* Kenardy JA, Dow MG, Johnston DW, Newman MG, Thomson A, Taylor CB. A comparison of delivery methods of cognitive-behavioral therapy for panic disorder: an international multicenter trial. *Journal of Consulting and Clinical Psychology* 2003;**71**(6):1068-75.

Dreessen 1994 *{published data only (unpublished sought but not used)}*

Arntz A, van den Hout M. Psychological treatments of panic disorder without agoraphobia: cognitive therapy versus applied relaxation. *Behaviour Research and Therapy* 1996;**34**(2):113-21.

Emmelkamp 1986 *{published data only (unpublished sought but not used)}*

Emmelkamp PM, Brilman E, Kuiper H, Mersch PP. The treatment of agoraphobia. A comparison of self-instructional training, rational emotive therapy, and exposure in vivo. *Behavior Modification* 1986;**10**:37-53.

Erickson 2003 *{published data only (unpublished sought but not used)}*

Erickson DH, Janeck AS, Tallman K. A cognitive-behavioral group for patients with various anxiety disorders. *Psychiatric Services* 2007;**58**(9):1205-11.

Erickson DH, Janeck AS, Tallman K. Group CBT for heterogeneous anxiety disorders. 34th Annual Conference of the British Association for Behavioural and Cognitive Psychotherapies. Warwick, 19-21 July 2006.

Erickson DH, Janeck AS, Tallman K. Group CBT for heterogeneous anxiety disorders: preliminary evidence for effectiveness. 31st Annual Conference of the British Association for Behavioural and Cognitive Psychotherapies. York, 16-19 July 2003.

Gloster 2010 *{published data only}*

Arolt V, Zwanziger P, Strohle A, Hamm A, Gerlach A, Kircher T, et al. The research network PANIC-NET: improving the treatment of panic disorder - From a better understanding of fear circuit mechanisms to more effective psychological treatment and routine care. *Psychotherapie Psychosomatik Medizinische Psychologie* 2009;**59**:124-31.

Emmrich A, Beesdo-Baum K, Gloster AT, Knappe S, Hfler M, Arolt V, et al. Depression does not affect the treatment outcome of CBT for panic and agoraphobia: results from a multicenter randomized trial. *Psychotherapy and Psychosomatics* 2012;**81**(3):161-72.

Gloster AT, Hauke C, Hofler M, Einsle F, Fydrich T, Hamm A, et al. Long-term stability of cognitive behavioral therapy effects for panic disorder with agoraphobia: a two-year follow-up study. *Behaviour Research and Therapy* 2013;**51**(12):830-9.

Gloster AT, Klotsche J, Gerlach AL, Hamm A, Strohle A, Gauggel S, et al. Timing matters: change depends on the stage of treatment in cognitive behavioral therapy for panic disorder with agoraphobia. *Journal of Consulting and Clinical Psychology* 2014;**82**(1):141-53.

Gloster AT, Wittchen HU, Einsle F, Hofler M, Lang T, Helbig-Lang S, et al. Mechanism of action in CBT (MAC): methods of a multi-center randomized controlled trial in 369 patients with panic disorder and agoraphobia. *European Archives of Psychiatry & Clinical Neuroscience* 2009;**259**(Suppl 2):S155-66.

Gloster AT, Wittchen HU, Einsle F, Lang T, Helbig-Lang S, Fydrich T, et al. Correction to Gloster et al. (2011). *Journal of Consulting and Clinical Psychology* 2011;**79**(5):652.

* Gloster AT, Wittchen HU, Einsle F, Lang T, Helbig-Lang S, Fydrich T, et al. Psychological treatment for panic disorder with agoraphobia: a randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology* 2011;**79**(3):406-20.

Kircher T, Arolt V, Jansen A, Pyka M, Reinhardt I, Kellermann T, et al. Effect of cognitive-behavioral therapy on neural correlates of fear conditioning in panic disorder. *Biological Psychiatry* 2013;**73**(1):93-101.

Lang T, Helbig-Lang S, Gloster AT, et al. Effects of therapist-guided exposure in CBT for panic disorder and agoraphobia [Effekte therapeutenbegleiteter versus patientengeleiteter Exposition bei Panikstörung mit Agoraphobie]. *Zeitschrift für klinische Psychologie und Psychotherapie* 2012;**41**(2):114-24.

Wittchen HU, Lang T. Improving cognitive behavioural therapy for panic by identifying the active ingredients and understanding the mechanisms of action: a multicentre study. http://www.isrctn.com/ISRCTN80046034 2007 (accessed 16 March 2015).

Goldstein 2000 *{published data only (unpublished sought but not used)}*

Goldstein AJ, De Beurs E, Chambless DL, Wilson KA. EMDR for panic disorder with agoraphobia: comparison with waiting list and credible attention-placebo control conditions. *Journal of Consulting and Clinical Psychology* 2000;**68**(6):947-56.

Gould 1993 *{published data only}*

Gould R, Clum GA, Shapiro D, Weaver T, Blalock J. Evidence for a self-help coping approach for treating panic disorder. Annual Meeting of the Southeastern Psychological Association. New Orleans, LA, 1991.

* Gould RA, Clum GA, Shapiro D. The use of bibliotherapy in the treatment of panic: a preliminary investigation. *Behavior Therapy* 1993;**24**:241-52.

Griegel 1995 *{published data only}*

Griegel LE. Breathing retraining in panic disorder: physiological mechanisms or perceived controllability. *Dissertation Abstracts International: Section B: The Sciences and Engineering* 1995;**55**(9):4120.

Hazen 1996 *{published data only (unpublished sought but not used)}*

Hazen AL, Walker JR, Eldridge GD. Anxiety sensitivity and treatment outcome in panic disorder. *Anxiety* 1996;**2**(1):34-9.

Walker JG, Eldridge A, Hazen, Rallo J, O'Riordan J, Frankel S. Evaluation of a self-help book and workbook for treatment of panic disorder used independently, in a self-help group, or a professional group [Thesis]. University of Manitoba 1996.

Hendriks 2010 *{published data only}*

Hendriks GJ. The efficacy of paroxetine and cognitive-behavioural therapy in the treatment of late-life panic disorder: a randomized controlled trial. http://www.trialregister.nl/ 2007 (accessed 16 March 2015). [NTR1144]

Hendriks GJ, Keijsers GPJ, Kampman M, Hoogduin CAL, Oude Voshaar RC. Predictors of outcome of pharmacological and psychological treatment of late-life panic disorder with

agoraphobia. *International Journal of Geriatric Psychiatry* 2012;**27**(2):146-50.

Hendriks GJ, Keijsers GPJ, Kampman M, Oude Voshaar RC, Verbraak M, Broekman TG, et al. A randomized controlled study of paroxetine and cognitive-behavioural therapy for late-life panic disorder. *Acta Psychiatrica Scandinavica* 2010;**122**(1):11-9.

Hoffart 1995 *{published data only}*

* Hoffart A. A comparison of cognitive and guided mastery therapy of agoraphobia. *Behaviour Research and Therapy* 1995;**33**(4):423-34.

Hoffart A. Cognitive and guided mastery therapy of agoraphobia: long-term outcome and mechanisms of change. *Cognitive Therapy and Research* 1998;**22**(3):195-207.

Hoffart A. Cognitive mediators of situational fear in agoraphobia. *Journal of Behavior Therapy and Experimental Psychiatry* 1995;**26**(4):313-20.

Hoffart A. Interpersonal problems among patients suffering from panic disorder with agoraphobia before and after treatment. *British Journal of Medical Psychology* 1997;**70**(2):149-57.

Hoffart A, Hedley LM. Personality traits among panic disorder with agoraphobia patients before and after symptom-focused treatment. *Journal of Anxiety Disorders* 1997;**11**(1):77-87.

Hoffart A, Hedley LM, Thornes K, Larsen SM, Friis S. Therapists' emotional reactions to patients as a mediator in cognitive behavioural treatment of panic disorder with agoraphobia. *Cognitive Behaviour Therapy* 2006;**35**(3):174-82.

Karekla 2004 *{published data only (unpublished sought but not used)}*

Karekla M. A Comparison Between Acceptance Enhanced Cognitive Behavioral and Panic Control Treatment for Panic Disorder [Doctoral dissertation]. University at Albany, State University of New York, 2004.

Karekla M. A comparison between acceptance-enhanced panic control treatment and panic control treatment for panic disorder. *Dissertation Abstracts International* 2005;**65**(9b):4835.

Levitt Jill T, Karekla M. Integrating acceptance and mindfulness with cognitive behavioral treatment for panic disorder. Acceptance and mindfulness-based approaches to anxiety: conceptualization and treatment. Springer US, 2005:165-88.

Klosko 1988 *{published data only}*

Klosko JS. Comparison of alprazolam and cognitive-behavior therapy in treatment of panic disorder. *Dissertation Abstracts International* 1988;**49**(5-b):1945.

Klosko JS, Barlow DH, Tassinari R, Cerny JA. A comparison of alprazolam and behavior therapy in treatment of panic disorder. *Journal of Consulting and Clinical Psychology* 1990;**58**(1):77-84.

Klosko JS, Barlow DH, Tassinari RB, Cerny JA. Comparison of alprazolam and cognitive-behavior therapy in the treatment of panic disorder. A preliminary report. Panic and Phobias: Treatment and Variables Affecting Course and Outcome. Springer-Verlag, 1988.

Korrelboom 2013 *{published data only}*

Korrelboom K, Peeters S, Blom S, Huijbrechts I. Competitive memory training (COMET) for panic and applied relaxation (AR) are equally effective in the treatment of panic in panic-disordered patients. *Journal of Contemporary Psychotherapy* 2014;**44**(3):183-90. [DOI: 10.1007/s10879-013-9259-3]

Lidren 1994 *{published data only}*

Lidren DM. The differential efficacy of self-help and group therapy in the treatment of panic disorder and agoraphobia: a treatment comparison study. *Dissertation Abstracts International* 1994;**54**(12-b):6465.

* Lidren DM, Watkins PL, Gould RA, Clum GA, Asterino M, Tulloch HL. A comparison of bibliotherapy and group therapy in the treatment of panic disorder. *Journal of Consulting and Clinical Psychology* 1994;**62**(4):865-9.

Malbos 2011 *{published and unpublished data}*

* Malbos E, Rapee RM, Kavakli M. A controlled study of agoraphobia and the independent effect of virtual reality exposure therapy. *Australian and New Zealand Journal of Psychiatry* 2013;**47**(2):160-8.

Malbos E, Rapee RM, Kavakli M. Isolating the effect of virtual reality based exposure therapy for agoraphobia: a comparative trial. *Studies in Health Technology and Informatics* 2011;**167**:45-50.

Marchione 1987 *{published data only}*

Marchione KE, Michelson L, Greenwald M, Dancu C. Cognitive behavioral treatment of agoraphobia. *Behaviour Research and Therapy* 1987;**25**(5):319-28.

Meulenbeek 2008 *{published and unpublished data}*

Meulenbeek P, Spinhoven P, Smit F, van Balkom A, Cuijpers P. Cognitive mediation of panic reduction during an early intervention for panic. *Acta Psychiatrica Scandinavica* 2010;**122**(1):20-9.

* Meulenbeek P, Willemse G, Smit F, van Balkom A, Spinhoven P, Cuijpers P. Early intervention in panic: pragmatic randomised controlled trial. *British Journal of Psychiatry* 2010;**196**:326-31.

Meulenbeek P, Willemse G, Smit F, van Balkom A, Spinhoven P, Cuijpers P. Early intervention in panic: pragmatic randomised controlled trial: Erratum. *British Journal of Psychiatry* 2010;**196**(6):497.

Meulenbeek P, Willemse G, Smit F, van Balkom A, Spinhoven P, Cuijpers P. Early intervention in panic: randomized controlled trial and cost-effectiveness analysis. *Trials* 2008;**9**:67.

Willemse G. Prevention of panic disorder: a randomised clinical trial adjoining cost-effectiveness study. http://www.isrctn.com/ISRCTN33407455 2005 (accessed 16 March 2015).

Meuret 2008 *{published data only (unpublished sought but not used)}*

Meuret AE, Wilhelm FH, Ritz T, Roth WT. Feedback of end-tidal pCO2 as a therapeutic approach for panic disorder. *Journal of Psychiatric Research* 2008;**42**(7):560-8.

Meuret 2010 *{published data only (unpublished sought but not used)}*

Meuret AE, Hofmann SG, Rosenfield D. Catastrophic appraisal and perceived control as moderators of treatment response in panic disorder. *International Journal of Cognitive Therapy* 2010;**3**(3):262-77.

* Meuret AE, Rosenfield D, Seidel A, Bhaskara L, Hofmann SG. Respiratory and cognitive mediators of treatment change in panic disorder: evidence for intervention specificity. *Journal of Consulting and Clinical Psychology* 2010;**78**(5):691-704.

Meuret AE, Seidel A, Rosenfield B, Hofmann SG, Rosenfield D. Does fear reactivity during exposure predict panic symptom reduction?. *Journal of Consulting and Clinical Psychology* 2012;**80**(5):777-85.

Meyerbroker 2011 *{published and unpublished data}*

Emmelkamp PMG. Virtual reality exposure therapy in agoraphobic participants. https://clinicaltrials.gov/ct2/show/NCT00734370 2008 (accessed 16 March 2015).

Meyerbroker K, Morina N, Kerkhof G, Emmelkamp PM. Virtual reality exposure treatment of agoraphobia: a comparison of computer automatic virtual environment and head-mounted display. *Studies in Health Technology & Informatics* 2011;**167**:51-6.

* Meyerbroker K, Morina N, Kerkhof G, Emmelkamp PMG. Virtual reality exposure therapy does not provide any additional value in agoraphobic patients: a randomized controlled trial. *Psychotherapy and Psychosomatics* 2013;**82**(3):170-6.

Meyerbroker K, Morina N, Kerkhof G, Emmelkamp PMG. Virtual reality exposure treatment of agoraphobia: a comparison of computer automatic virtual environment and head-mounted display. *Annual Review of CyberTherapy and Telemedicine* 2011;**9**(1):41-5.

Milrod 2006a *{published and unpublished data}*

Klass ET, Milrod BL, Leon AC, Kay SJ, Schwalberg M, Li C, et al. Does interpersonal loss preceding panic disorder onset moderate response to psychotherapy? An exploratory study. *Journal of Clinical Psychiatry Diseases of the Nervous System* 2009;**70**(3):406-11.

Milrod B. Outcome studies in psychodynamic psychotherapy for panic disorder. 37th International Meeting of the Society for Psychotherapy Research; 2006 21-24 June. Edinburgh, 2006:103-4.

Milrod B. Randomized controlled trial of psychodynamic psychotherapy vs. applied relaxation for panic disorder. https://clinicaltrials.gov/ct2/show/NCT00128388 2005 (accessed 16 March 2015). [NCT00128388]

* Milrod B, Leon AC, Busch F, Rudden M, Schwalberg M, Clarkin J, et al. A randomized controlled clinical trial of psychoanalytic psychotherapy for panic disorder [published erratum appears in The American Journal of Psychiatry 2009;164(7):1123]. *American Journal of Psychiatry* 2007;**164**(2):265-72.

Milrod B, Leon AC, Busch F, Rudden M, Schwalberg M, Clarkin J, et al. A randomized controlled clinical trial of psychoanalytic psychotherapy for panic disorder: Erratum. *The American Journal of Psychiatry* 2007;**164**(3):529.

Muncy 1991 *{published data only}*

Muncy SM. Panic: a comparison of four treatment methods. *Dissertation Abstracts International* 1991;**51**(12b, pt 1):6115.

Ost 1993 *{published data only (unpublished sought but not used)}*

Ost LG, Westling BE, Hellstrom K. Applied relaxation, exposure in vivo and cognitive methods in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy* 1993;**31**(4):383-94.

Ost 1995 *{published data only}*

Ost LG, Westling BE. Applied relaxation vs cognitive behavior therapy in the treatment of panic disorder. *Behaviour Research and Therapy* 1995;**33**(2):145-58.

Ost 2004 *{published data only (unpublished sought but not used)}*

Ost LG, Thulin U, Ramnero J. Cognitive behavior therapy vs exposure in vivo in the treatment of panic disorder with agoraphobia. *Behaviour Research and Therapy* 2004;**42**(10):1105-27.

Ramnero J, Ost LG. Panic and avoidance in panic disorder with agoraphobia: clinical relevance of change in different aspects of the disorder. *Journal of Behavior Therapy and Experimental Psychiatry* 2007;**38**(1):29-39.

Petterson 1996 *{published data only}*

Petterson K, Cesare S. Panic disorder: a cognitive-behavioural approach to treatment. *Counselling Psychology Quarterly* 1996;**9**(2):191-201.

Rees 1999 *{published data only}*

Rees CS, Richards JC, Smith LM. The efficacy of information-giving in cognitive-behavioural treatment for panic disorder. *Behaviour Change* 1999;**16**(3):175-81.

Reinecke 2013 *{published data only}*

Reinecke A, Waldenmaier L, Cooper MJ, Harmer CJ. Changes in automatic threat processing precede and predict clinical changes with exposure-based cognitive-behavior therapy for panic disorder. *Biological Psychiatry* 2013;**73**(11):1064-70.

Salkovskis 1999 *{published data only}*

Salkovskis PM, Clark DM, Hackmann A, Wells A, Gelder MG. An experimental investigation of the role of safety-seeking behaviours in the maintenance of panic disorder with agoraphobia. *Behaviour Research and Therapy* 1999;**37**(6):559-74.

Salkovskis PM, Hackmann A, Wells A, Gelder MG, Clark DM. Belief disconfirmation versus habituation approaches to situational exposure in panic disorder with agoraphobia: a pilot study. *Behaviour Research and Therapy* 2007;**45**(5):877-85.

Schmidt 1997a *{published data only}*

Schmidt NB, Trakowski JH, Staab JP. Extinction of panicogenic effects of a 35% CO2 challenge in patients with panic disorder. *Journal of Abnormal Psychology* 1997;**106**(4):630-8.

Schmidt 1997b *{published data only (unpublished sought but not used)}*

* Schmidt NB, Staab JP, Trakowski JH Jr, Sammons M. Efficacy of a brief psychosocial treatment for panic disorder in an active duty sample: implications for military readiness. *Military Medicine* 1997;**162**(2):123-9.

Schmidt NB, Woolaway-Bickel K, Trakowski J, Santiago H, Storey J, Koselka M, et al. Dismantling cognitive-behavioral treatment for panic disorder: questioning the utility of breathing retraining. *Journal of Consulting and Clinical Psychology* 2000;**68**(3):417-24.

Scott 1995 *{published data only (unpublished sought but not used)}*

Scott MJ, Stradling SG, Greenfield TA. The efficacy of brief group cognitive therapy programmes for anxiety and depression, and the relevance of a personality disorder diagnosis. World Congress of Behavioural and Cognitive Therapies; 1995 July 11-15. Copenhagen, 1995.

Sharp 2004 *{published data only}*

Sharp DM, Power KG, Swanson V. A comparison of the efficacy and acceptability of group versus individual cognitive behaviour therapy in the treatment of panic disorder and agoraphobia in primary care. *Clinical Psychology and Psychotherapy* 2004;**11**(2):73-82.

Shear 1994 *{published data only}*

Shear MK, Leon AC, Portera L, Klosko J, Cloitre M. Psychoeducation /reflective listening compared to cognitive-behavioural treatment in panic disorder. New Research Program and Abstracts - 144th Annual Meeting of the American Psychiatric Association; 1991 May 11-16. New Orleans, Louisiana, US, 1991.

* Shear MK, Pilkonis PA, Cloitre M, Leon AC. Cognitive behavioral treatment compared with nonprescriptive treatment of panic disorder. *Archives of General Psychiatry* 1994;**51**(5):395-401.

Taylor 1982 *{published data only (unpublished sought but not used)}*

Taylor CB, Kenigsberg ML, Robinson JM. A controlled comparison of relaxation and diazepam in panic disorder. *Journal of Clinical Psychiatry* 1982;**43**(10):423-5.

Telch 1993 *{published data only}*

Telch MJ, Lucas JA, Schmidt NB, Hanna HH, LaNae Jaimez T, Lucas RA. Group cognitive-behavioral treatment of panic disorder. *Behaviour Research and Therapy* 1993;**31**(3):279-87.

Tyrer 1988 *{published data only (unpublished sought but not used)}*

Kingdon D, Tyrer P, Seivewright N, Ferguson B, Murphy S. The Nottingham Study of Neurotic Disorder: influence of cognitive therapists on outcome. *British Journal of Psychiatry* 1996;**169**:93-7.

Klein DF. Nottingham study of neurotic disorder. *Lancet* 1988;**2**:1015.

Knerer G, Byford S, Johnson T, Seivewright H, Tyrer P. The Nottingham study of neurotic disorder: predictors of 12 year costs. *Acta Psychiatrica Scandinavica* 2005;**112**(3):224-32.

Seivewright H, Tyrer P, Johnson T. Change in personality status in neurotic disorders. *Lancet* 2002;**359**:2253-4.

Seivewright H, Tyrer P, Johnson T. Prediction of outcome in neurotic disorder: a 5-year prospective study. *Psychological Medicine* 1998;**28**(5):1149-57.

Seivewright N, Tyrer P, Ferguson B, Murphy S, Johnson T. Longitudinal study of the influence of life events and personality status on diagnostic change in three neurotic disorders. *Depression & Anxiety* 2000;**11**(3):105-13.

Tyrer P, Seivewright H, Ferguson B, Johnson T. Cold calling in psychiatric follow up studies: is it justified?. *Journal of Medical Ethics* 2003;**29**(4):238-42.

Tyrer P, Seivewright H, Johnson T. The Nottingham Study of Neurotic Disorder: Predictors of 12-year outcome of dysthymic, panic and generalized anxiety disorder. *Psychological Medicine* 2004;**34**(8):1385-94.

Tyrer P, Seivewright H, Simmonds S, Johnson T. Prospective studies of cothymia (mixed anxiety-depression): How do they inform clinical practice?. *European Archives of Psychiatry & Clinical Neuroscience* 2001;**251**(Suppl 2):1153-6.

Tyrer P, Seivewright N, Ferguson B, Murphy S, Darling C, Brothwell J, et al. The Nottingham Study of Neurotic Disorder: relationship between personality status and symptoms. *Psychological Medicine* 1990;**20**(2):423-31.

Tyrer P, Seivewright N, Ferguson B, Murphy S, Johnson AL. The Nottingham Study of Neurotic Disorder: effect of personality status on response to drug treatment, cognitive therapy and self-help over two years. *British Journal of Psychiatry* 1993;**162**:219-26.

Tyrer P, Seivewright N, Ferguson B, Tyrer J. The general neurotic syndrome: a coaxial diagnosis of anxiety, depression and personality disorder. *Acta Psychiatrica Scandinavica* 1992;**85**(3):201-6.

* Tyrer P, Seivewright N, Murphy S, Ferguson B, Kingdon D, Barczak P, et al. The Nottingham Study of Neurotic Disorder: comparison of drug and psychological treatments. *Lancet* 1988;**2**:235-40.

Tyrer P, Seivewright N, Seivewright H. Long-term outcome of hypochondriacal personality disorder. *Journal of Psychosomatic Research* 1999;**46**(2):177-85.

Williams 1996 *{published data only (unpublished sought but not used)}*

Williams SL, Falbo J. Cognitive and performance-based treatments for panic attacks in people with varying degrees of agoraphobic disability. *Behaviour Research and Therapy* 1996;**34**(3):253-64.

Wollburg 2011 *{published data only}*

* Kim S, Wollburg E, Roth WT. Opposing breathing therapies for panic disorder: a randomized controlled trial of lowering vs raising end-tidal PCO2. *Journal of Clinical Psychiatry* 2012;**73**(7):931-9.

Roth WT. Breathing regulation training for individuals with panic disorder. https://clinicaltrials.gov/ct2/show/NCT00183521 2005 (accessed 16 March 2015).

Wollburg E, Roth WT, Kim S. Effects of breathing training on voluntary hypo- and hyperventilation in patients with panic disorder and episodic anxiety. *Applied Psychophysiology & Biofeedback* 2011;**36**:81-91.

References to studies excluded from this review

Andersson 2011 *{published and unpublished data}*

Andersson G. Internet-administrated treatment of anxiety symptoms for young adults. https://clinicaltrials.gov/ct2/show/NCT01402258 2011 (accessed 16 March 2015).

Barlow 2000a *{published data only}*

* Barlow DH, Gorman JM, Shear MK, Woods SW. Cognitive-behavioral therapy, imipramine, or their combination for panic disorder. a randomized controlled trial. *JAMA* 2000;**283**(19):2529-36.

Barlow DH, Gorman JM, Shear MK, Woods SW. ERRATUM: Cognitive-behavioral therapy, imipramine, or their combination for panic disorder. a randomized controlled trial. *JAMA* 2000;**284**(19):2450.

Barlow DH, Gorman JM, Shear MK, Woods SW. ERRATUM: Cognitive-behavioral therapy, imipramine, or their combination for panic disorder. a randomized controlled trial. *JAMA* 2001;**284**(20):2597.

Shear MK, Houck P, Greeno C, Masters S. Emotion-focused psychotherapy for patients with panic disorder. *American Journal of Psychiatry* 2001;**158**(12):1993-8.

Bélanger 2006 *{published and unpublished data}*

Bélanger C, Vaillancourt L, Dulude D, Archambault M, Rochfort J, Pecknold J, et al. The differential effect of two behavioral therapies on attentional processes of subjects presenting panic disorder [L'effet différentiel de deux thérapies comportementales sur les processus attentionnels de sujets présentant un trouble panique]. *Revue Francophone de Clinique Comportamentale et Cognitive* 2008;**13**(2):11-23.

Benecke 2014 *{published data only}*

Benecke C. Psychoanalytic therapy (PT) and cognitive-behavioral therapy (CBT) in outpatients with anxiety (panic disorder/agoraphobia) and comorbid personality disorders: a multicenter prospective randomized superiority trial. http://www.isrctn.com/ISRCTN12449681 2014 (accessed 16 March 2015).

Borden 1986 *{published data only}*

Clum GA, Watkins PL, Borden JW, Broyles S, Hayes J. A comparison of guided imaginal coping and imaginal exposure in the treatment of panic disorder. *Journal of Rational-Emotive & Cognitive-Behavior Therapy* 1993;**11**(4):179.

Elsesser 2002 *{published and unpublished data}*

Elsesser K, Mosch A, Sartory G. Brief psychological treatment for the relief of panic disorder. *Behavioural and Cognitive Psychotherapy* 2002;**30**(4):423-30.

Fava 1997 *{published data only (unpublished sought but not used)}*

Fava GA, Savron G, Zielezny M, Grandi S, Rafanelli C, Conti S. Overcoming resistance to exposure in panic disorder with agoraphobia. *Acta Psychiatrica Scandinavica* 1997;**95**(4):306-12.

Gloster 2010a *{published data only (unpublished sought but not used)}*

Gloster AT. Acceptance and commitment therapy (ACT) for treatment-resistant panic disorder with agoraphobia. http://www.isrctn.com/ISRCTN12042066 2010 (accessed 16 March 2015).

Ito 2001 *{published data only}*

Ito LM, De Araujo LA, Tess VLC, De Barros-Neto TP, Asbahr FR, Marks I. Self-exposure therapy for panic disorder with agoraphobia. *British Journal of Psychiatry* 2001;**178**(4):331-6.

Michelson 1996 *{published data only}*

Michelson LK, Marchione KE, Greenwald MT, Testa S, Marchione NJ. A comparative outcome and follow-up investigation of panic disorder with agoraphobia: the relative and combined efficacy of cognitive therapy, relaxation training, and therapist-assisted exposure. *Journal of Anxiety Disorders* 1996;**10**(5):297-330.

Murphy MT, Michelson LK, Marchione K, Marchione N, Testa S. The role of self-directed in vivo exposure in combination with cognitive therapy, relaxation training, or therapist-assisted exposure in the treatment of panic disorder with agoraphobia. *Journal of Anxiety Disorders* 1998;**12**(2):117-38.

Teusch 1996 *{published data only (unpublished sought but not used)}*

Teusch L, Bohme H, Gastpar M. The benefit of an insight-oriented and experiential approach on panic and agoraphobia symptoms. Results of a controlled comparison of client-centered therapy alone and in combination with behavioral exposure. *Psychotherapy & Psychosomatics* 1997;**66**:293-301.

Zane 1993 *{published data only (unpublished sought but not used)}*

Zane G, Williams SL. Performance-related anxiety in agoraphobia: treatment procedures and cognitive mechanisms of change. *Behavior Therapy* 1993;**196**:625-43.

References to studies awaiting assessment

Bressi 2010a *{published data only (unpublished sought but not used)}*

Bressi C, Ciabatti M, Nocito EP, Catenacci E, Porcellana, M, Invernizzi G, et al. A preliminary longitudinal study on the importance of early intervention in panic disorder [conference abstract]. Journal of Psychosomatic Research [abstracts from the 19th European Congress of Psychiatry, EPA 2011 Vienna Austria, 12-15 Mar]. 2011:610.

Foley 2006 *{published data only (unpublished sought but not used)}*

Foley D, Baille A, Renner P. 29th Australian Association for Cognitive and Behaviour Therapy Annual Conference. Manly, 18-23 October 2006:43.

Franklin 1990 *{published data only}*

Franklin JA. Behavioural therapy for panic disorder. *International Journal of Neuroscience* 1990;**51**:138.

Irgens 2009 *{published data only (unpublished sought but not used)}*

Irgens AC. Thought field therapy and cognitive therapy for agoraphobia. https://clinicaltrials.gov/ct2/show/NCT00932919 2009 (accessed 16 March 2015).

Margraf 1991 *{published data only (unpublished sought but not used)}*

Margraf J, Schneider S. Outcome and active ingredients of cognitive-behavioural treatments for panic disorder. 25th Conference of the Association for Advancement of Behavior Therapy. New York, 1991.

Milrod 2006b *{published data only (unpublished sought but not used)}*

Milrod B, Leon AC, Fishman B, Barber JP, Chambless D. Dynamic treatment vs. CBT for panic disorder. https://clinicaltrials.gov/ct2/show/NCT00353470 2006 (accessed 16 March 2015).

Richards 1997 *{published data only (unpublished sought but not used)}*

Richards JB, Bickley N, Rees C, Beros P. An investigation of the mechanisms of change in the cognitive behavioural treatment of panic disorder. *Health Perspectives: Research Policy and Practice* 1997;**1**(1):35-44.

Roache 1998 *{published data only (unpublished sought but not used)}*

Roache JD, Oswald LM, Stanley MA, Creason DR, Shah NN. Effects of psychotherapy-induced anxiety reduction on alprazolam self-medication behavior. CPDD 1998 Annual Meeting. Scottsdale, 1998:121.

Strauss 1997 *{published data only (unpublished sought but not used)}*

Strauss WH, Klieser E. Combination of pharmaco- and psychotherapy in the treatment of panic disorder. *European Neuropsychopharmachology* 1996;**6**(Suppl 3):206 (O-24-6).

Vincelli 2003 *{published data only (unpublished sought but not used)}*

Vincelli F, Anolli L, Bouchard S, Wiederhold BK, Zurloni V, Riva G. Experiential cognitive therapy in the treatment of panic disorders with agoraphobia: a controlled study. *Cyberpsychology and Behavior* 2003;**6**(3):321-8.

Vincelli 2004 *{published data only (unpublished sought but not used)}*

Vincelli F, Molinari E, Riva G. CBT Virtual Reality Assisted for the Treatment of Panic Disorders with Agoraphobia: A Controlled Study. 35th International Meeting of the Society for Psychotherapy Research. Rome, 16-19 June 2004:35.

References to ongoing studies

Barlow 2010 *{published data only (unpublished sought but not used)}*

Barlow DH. Efficacy evaluation of a unified transdiagnostic treatment for anxiety disorders. https://clinicaltrials.gov/ct2/show/NCT01243606 2010 (accessed 16 March 2015).

Caspi 2012 *{published data only (unpublished sought but not used)}*

Caspi A. The assessment and treatment of balance impairment using virtual reality (VR) in panic disorder patients. https://clinicaltrials.gov/ct2/show/NCT01677429 2012 (accessed 16 March 2015).

Gensichen 2012 *{published data only (unpublished sought but not used)}*

Gensichen J. Evaluation of a practice team-supported, self-managed exposure training for patients with panic disorder and agoraphobia in primary care [Jena-PARADISE (Patient Activation foR Anxiety DISordErs)]. http://www.isrctn.com/ISRCTN64669297 2012 (accessed 16 March 2015).

Gensichen J, Hiller TS, Breitbart J, Teismann T, Brettschneider C, Schumacher U, Piwtorak A, König HH, Hoyer H, Schneider N, Schelle M, Blank W, Thiel P, Wensing M, Margraf J. Evaluation of a practice team-supported exposure training for patients with panic disorder with or without agoraphobia in primary care - study protocol of a cluster randomised controlled superiority trial [2014]. *Trials* 2014;**15**:112.

Sandell 2012 *{published data only}*

Sandell R. Psychotherapy outcome and self-selection effects in panic disorder. https://clinicaltrials.gov/ct2/show/NCT01606592 2012 (accessed 16 March 2015).

Teismann 2012 *{published data only}*

Teismann T. Cognitive behavior therapy vs exposure in vivo in the treatment of panic disorder with agoraphobia. https://clinicaltrials.gov/ct2/show/NCT01680237 2012 (accessed 16 March 2015).

Additional references

Altman 1996

Altman DG, Bland MJ. Detecting skewness for summary information. *BMJ* 1996;**313**:1200.

APA 1980

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3rd Edition. American Psychiatric Association, 1980.

APA 1987

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 3rd Edition. American Psychiatric Association, 1987.

APA 2000

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 4th Edition. Washington DC: APA, 2000.

APA 2013

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. 5th Edition. Arlington, VA: American Psychiatric Publishing, 2013.

Barlow 2000b

Barlow DH, Craske MG. Mastery of Your Anxiety and Panic: MAP-3. New York: Graywind Publications, 2000.

Bernstein 1973

Bernstein DA, Borkovec TD. Progressive Relaxation Training: A Manual for the Helping Professions. Champaign, IL: Research Press, 1973.

Bijl 1998

Bijl RV, Ravelli A, van Zessen G. Prevalence of psychiatric disorder in the general population: results of The Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Social Psychiatry and Psychiatric Epidemiology* 1998;**33**:587-95.

Brooks 1998

Brooks SP, Gelman A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 1998;**7**:434–55.

Bucher 1997

Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology* 1997;**50**(6):683-91.

Busch 1996

Busch F, Milrod B, Cooper A, Shapiro T. Psychodynamic approaches to panic disorder. *Journal of Psychotherapy Practice and Research* 1996;**5**(1):72-83.

Caldwell 2005

Caldwell DM, Ades AE, Higgins JP. Simultaneous comparison of multiple treatments: combining direct and indirect evidence. *BMJ* 2005;**331**(7521):897-900.

Carlbring 2006

Carlbring P, Bohman S, Brunt S, Buhrman M, Westling BE, Ekselius L, et al. Remote treatment of panic disorder: a randomized trial of internet-based cognitive behavior therapy supplemented with telephone calls. *American Journal of Psychiatry* 2006;**163**(12):2119-25.

Chaimani 2013

Chaimani A, Higgins JPT, Mavridis D, Spyridonos P, Salanti G. Graphical tools for network meta-analysis in STATA. *PLoS ONE* 2013;**8**(10):e76654. [DOI: 10.1371/journal.pone.0076654]

Clark 1985

Clark DM, Salkovskis PM, Chakley J. Respiratory control as a treatment for panic attacks. *Journal of Behavior Therapy and Experimental Psychiatry* 1985;**16**:23-30.

Clark 1986a

Clark DM. A cognitive approach to panic. *Behaviour Research and Therapy* 1986;**24**(4):461-70.

Clark 1986b

Clark DA, Salkovskis PM. Cognitive Treatment of Panic: Therapist's Manual. Oxford, UK: Department of Psychiatry, University of Oxford, 1986.

Cochrane Comparing Multiple Interventions Group

Cochrane Comparing Multiple Interventions Group. [Stream 3: summary of findings tables and GRADE assessments for network meta-analysis]. http://methods.cochrane.org/cmi/sites/methods.cochrane.org.cmi/files/uploads/SoFTs%20and%20GRADE%20for%20NMA.pdf (accessed 11 April 2016).

Cohen 1960

Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement* 1960;**20**:37-46. [DOI: 10.1177/001316446002000104]

Dannon 2002

Dannon PN, Iancu I, Grunhaus L. Psychoeducation in panic disorder patients: effect of a self-information booklet in a randomized, masked-rater study. *Depression and Anxiety* 2002;**16**(2):71-6.

Eaton 1994

Eaton WW, Kessler RC, Wittchen H-U, Magee WJ. Panic and panic disorder in the United States. *American Journal of Psychiatry* 1994;**151**:413-20.

Elbourne 2002

Elbourne DR, Altman DG, Higgins JP, Curtin F, Worthington HV, Vail A. Meta-analyses involving cross-over trials: methodological issues. *International Journal of Epidemiology* 2002;**31**:140-9. [DOI: 10.1093/ije/31.1.140]

Feighner 1972

Feighner JP, Robins E, Guze SB, Woodruff RA, Winokur G, Munoz R. Diagnostic criteria for use in psychiatric research. *Archives of General Psychiatry* 1972;**26**:57-63.

Furukawa 2005

Furukawa TA, Cipriani A, Barbui C, Brambilla P, Watanabe N. Imputing response rates from means and standard deviations in meta-analysis. *International Clinical Psychopharmacology* 2005;**20**:49-52.

Furukawa 2006

Furukawa TA, Barbui C, Cipriani A, Brambilla P, Watanabe N. Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology* 2000;**59**:7-10.

Furukawa 2007

Furukawa TA, Watanabe N, Churchill R. Combined psychotherapy plus antidepressants for panic disorder with or without agoraphobia. *Cochrane Database of Systematic Reviews* 2007, Issue 1. [DOI: 10.1002/14651858.CD004364.pub2]

Glenny 2005

Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, et al. Indirect comparisons of competing interventions. *Health Technology Assessment* 2005;**9**(26):1-134.

Gloster 2011

Gloster AT, Wittchen HU, Einsle F, Lang T, Helbig-Lang S, Fydrich T, et al. Psychological treatment for panic disorder with agoraphobia: a randomized controlled trial to examine the role of therapist-guided exposure in situ in CBT. *Journal of Consulting and Clinical Psychology* 2011;**79**(3):406-20.

Goisman 1995

Goisman RM, Warshaw MG, Steketee GS, Fierman EJ, Rogers MP, Goldenberg I, et al. DSM-IV and the disappearance of agoraphobia without a history of panic disorder: new data on a controversial diagnosis. *American Journal of Psychiatry* 1995;**152**:1438-43.

Gorman 2000

Gorman JM, Kent JM, Sullivan GM, Coplan JD. Neuroanatomical hypothesis of panic disorder, revised. *American Journal of Psychiatry* 2000;**157**:493-505.

Grant 2006

Grant BF, Hasin DS, Stinson FS, Dawson DA, Goldstein RB, Smith S, et al. The epidemiology of DSM-IV panic disorder and agoraphobia in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Journal of Clinical Psychiatry* 2006;**67**:363-74.

Greenberg 2004

Greenberg LS. Emotion–focused therapy. *Clinical Psychology and Psychotherapy* 2004;**11**:3-16.

Guaiana 2013

Guaiana G, Barbui C, Chiodo D, Cipriani A, Davies SJC, Imai H, et al. Azapirones versus placebo for panic disorder in adults. *Cochrane Database of Systematic Reviews* 2013, Issue 11. [DOI: 10.1002/14651858.CD010828]

Guaiana 2013a

Guaiana G, Barbui C, Chiodo D, Cipriani A, Davies SJC, Koesters M. Antidepressants versus placebo for panic disorder in adults. *Cochrane Database of Systematic Reviews* 2013, Issue 7. [DOI: 10.1002/14651858.CD010676]

Guaiana 2013b

Guaiana G, Barbui C, Chiodo D, Cipriani A, Davies SJC, Koesters M. Benzodiazepines versus placebo for panic disorder in adults. *Cochrane Database of Systematic Reviews* 2013, Issue 7. [DOI: 10.1002/14651858.CD010677]

Guy 1976

Guy W. Clinical Global Impressions. In: ECDEU Assessment Manual for Psychopharmacology, revised (DHEW Publ No ADM 76-338). National Institute of Mental Health, 1976:218–222.

Hayes 1999

Hayes SC, Strosahl KD, Wilson KG. Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change. New York, NY: Guilford Press, 1999.

Hayes 2004

Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy* 2004;**35**:639-65.

Higgins 2003

Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003;**327**(7414):557-60.

Higgins 2011

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Higgins 2012

Higgins JPT, Jackson D, Barrett JK, Lu G, Ades AE, White IR. Consistency and inconsistency in network meta-analysis: concepts and models for multi-arm studies. *Research Synthesis Methods* 2012;**3**(2):98-110.

Hofmann 2008

Hofmann SG, Smits JAJ. Cognitive-behavioral therapy for adult anxiety disorders: a meta-analysis of randomized placebo-controlled trials. *Journal of Clinical Psychiatry* 2008;**69**(4):621-32.

Kessler 2006

Kessler RC, Chiu WT, Jin R, Ruscio AM, Shear K, Walters EE. The epidemiology of panic attacks, panic disorder, and agoraphobia in the National Comorbidity Survey Replication. *Archives of General Psychiatry* 2006;**63**:415-24.

Kim 2009

Kim YW, Lee SH, Choi TK, Suh SY, Kim B, Kim CM, et al. Effectiveness of mindfulness-based cognitive therapy as an adjuvant to pharmacotherapy in patients with panic disorder or generalized anxiety disorder. *Depression and Anxiety* 2009;**26**(7):601-6.

King 2008

King M, Nazareth I, Levy G, Walker C, Morris R, Weich S, et al. Prevalence of common mental disorders in general practice attendees across Europe. *British Journal of Psychiatry* 2008;**192**:362-7.

Klein 1983

Klein DF, Zitrin CM, Woerner MG, Ross DC. Treatment of phobias. II. Behavior therapy and supportive psychotherapy: are there any specific ingredients?. *Archives of General Psychiatry* 1983;**40**(2):139-45.

Klein 1993

Klein DF. False suffocation alarms, spontaneous panics, and related conditions. An integrative hypothesis. *Archives of General Psychiatry* 1993;**50**:306–17.

Lee 2007

Lee SH, Ahn SC, Lee YJ, Choi TK, Yook KH, Suh SY. Effectiveness of a meditation-based stress management program as an adjunct to pharmacotherapy in patients with anxiety disorder. *Journal of Psychosomatic Research* 2007;**62**:189–95.

Ley 1985

Ley RA. Blood, breath and fears: a hyperventilation theory of panic attacks and agoraphobia. *Clinical Psychology Review* 1985;**5**:271–85.

Lu 2004

Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. *Statistics in Medicine* 2004;**30**(20):3105-24.

Ludwig 2008

Ludwig DS, Kabat-Zinn J. Mindfulness in medicine. *JAMA* 2008;**300**(11):1350-2.

Lumley 2002

Lumley T. Network meta-analysis for indirect treatment comparisons. *Statistics in Medicine* 2002;**21**(16):2313-24.

Marks 1979

Marks IM, Mathews AM. Brief standard self-rating for phobic patients. *Behavior Research and Therapy* 1979;**17**:263-7.

Marks 1981

Marks IM. Cure and Care of Neuroses: Theory and Practice of Behavioural Psychotherapy. New York: Wiley, 1981.

Meuret 2010a

Meuret AE, Rosenfield D, Seidel A, Bhaskara L, Hofmann SG. Respiratory and cognitive mediators of treatment change in panic disorder: evidence for intervention specificity. *Journal of Consulting and Clinical Psychology* 2010;**78**:691-704.

Meuret 2010b

Meuret AE, Ritz T. Hyperventilation in panic disorder and asthma: empirical evidence and clinical strategies. *International Journal of Psychophysiology* 2010;**78**(1):68-79.

Meuret 2012

Meuret AE, Wolitzky-Taylor KB, Twohig MP, Craske MG. Coping skills and exposure therapy in panic disorder and agoraphobia: lates *Behavior Therapy* 2012;**43**(2):271-84.

Milrod 1997

Milrod B, Busch F, Cooper A, Shapiro T. Manual of Panic-Focused Psychodynamic Psychotherapy. Arlington: American Psychiatric Publishing, 1997.

Milrod 2007

Milrod B, Leon AC, Busch F, Rudden M, Schwalberg M, Clarkin J, et al. A randomized controlled clinical trial of psychoanalytic psychotherapy for panic disorder. *American Journal of Psychiatry* 2007;**164**(2):265-72.

Mitte 2005

Mitte K. A meta-analysis of the efficacy of psycho- and pharmacotherapy in panic disorder with and without agoraphobia. *Journal of Affective Disorders* 2005;**88**(1):27-45.

NICE 2011

[CG113] Generalised anxiety disorder and panic disorder (with or without agoraphobia) in adults: management in primary, secondary and community care. NICE clinical guidelines 2011.

Nordin 2010

Nordin S, Carlbring P, Cuijpers P, Andersson G. Expanding the limits of bibliotherapy for panic disorder: randomized trial of self-help without support but with a clear deadline. *Behavior Therapy* 2010;**41**(3):267-76.

Ost 1987

Ost LG. Applied relaxation: description of a coping technique and review of controlled studies. *Behaviour Research and Therapy* 1987;**25**(5):397-409.

Ost 1988

Ost LG. Applied relaxation vs progressive relaxation in the treatment of panic disorder. *Behaviour Research and Therapy* 1988;**26**(1):13-22.

Roemer 2008

Roemer L, Erisman SM, Orsillo SM. Mindfulness and acceptance-based treatments for anxiety disorders. In: Antony MM, Stein MB editor(s). Oxford Handbook of Anxiety and Related Disorders. Oxford: Oxford University Press, 2008:476-87.

Rogers 1980

Rogers CR. A Way of Being. Boston: Houghton Mifflin, 1980.

Salanti 2008

Salanti G, Higgins JPT, Ades AE, Ioannidis JPA. Evaluation of networks of randomized trials. *Statistical Methods in Medical Research* 2008;**17**:279. [DOI: 10.1177/0962280207080643]

Salanti 2009

Salanti G, Marinho V, Higgins JP. A case study of multiple-treatments meta-analysis demonstrates that covariates should be considered. *Journal of Clinical Epidemiology* 2009;**62**(8):857-64.

Salanti 2011

Salanti G, Ades AE, Ioannidis JPA. Graphical methods and numerical summaries for presenting results from multiple-treatment meta-analysis: an overview and tutorial. *Journal of Clinical Epidemiology* 2011;**64**:163-71.

Salanti 2012

Salanti G. Indirect and mixed-treatment comparison, network, or multiple-treatments meta-analysis: many names, many benefits, many concerns for the next generation evidence synthesis tool. *Research Synthesis Methods* 2012;**3**(2):80-97.

Salanti 2014

Salanti G, Del Giovane C, Chaimani A, Caldwell DM, Higgins JP. Evaluating the quality of evidence from a network meta-analysis. *PLoS One* 2014;**9**(7):e99682.

Salkovskis 1991

Salkovskis PM, Clark DM, Hackmann A. Treatment of panic attacks using cognitive therapy without exposure or breathing retraining. *Behaviour Research and Therapy* 1991;**29**:161-6.

Schmidt 1994

Schmidt NB. Elimination of Safety Maneuvers: Cognitive Behavioral Treatment for Panic Disorder with Agoraphobia. Bethesda: USUHS, 1994.

Schulz 1995

Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**(5):408–12.

Shear 1997

Shear MK, Brown TA, Barlow DH, Money R, Sholomskas DE, Woods SW, et al. Multicenter collaborative panic disorder severity scale. *American Journal of Psychiatry* 1997;**154**(11):1571-5.

Sorby 1991

Sorby NG, Reavley W, Huber JW. Self help programme for anxiety in general practice: controlled trial of an anxiety management booklet. *British Journal of General Practice* 1991;**41**(351):417-20.

Spineli 2013

Spineli LM, Higgins JPT, Cipriani A, Leucht S, Salanti G. Evaluating the impact of imputations for missing participant outcome data in a network meta-analysis. *Clinical Trials* 2013;**10**(3):378-88. [DOI: 10.1177/1740774512470317]

Spitzer 1978

Spitzer RL, Robins E. Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry* 1978;**35**(6):773-82.

Starcevic 2009

Starcevic V. Anxiety Disorders in Adults: a Clinical Guide. Oxford University Press, 2009.

Sterne 2000

Sterne JA, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. *Journal of Clinical Epidemiology* 2000;**53**(11):1119-29.

Sánchez-Meca 2010

Sánchez-Meca J, Rosa-Alcázar AI, Marín-Martínez F, Gómez-Conesa A. Psychological treatment of panic disorder with or without agoraphobia: a meta-analysis. *Clinical Psychology Review* 2010;**30**(1):37-50.

Telch 1990

Telch MJ, Schmidt NB. Cognitive–behavioral Treatment for Panic Disorder and Agoraphobia: Panic Inoculation Treatment Manual. Unpublished manuscript, 1990.

Turner 2012

Turner RM, Davey J, Clarke MJ, Thompson SG, Higgins JP. Predicting the extent of heterogeneity in meta-analysis, using empirical data from the Cochrane Database of Systematic Reviews. *International Journal of Epidemiology* 2012;**41**(3):818-27.

Van den Hout 1994

Van den Hout M, Arntz A, Hoekstra R. Exposure reduced agoraphobia but not panic, and cognitive therapy reduced panic but not agoraphobia. *Behaviour Research and Therapy* 1994;**32**:447-51.

Vøllestad 2012

Vøllestad J, Nielsen MB, Nielsen GH. Mindfulness- and acceptance-based interventions for anxiety disorders: a systematic review and meta-analysis. *British Journal of Clinical Psychology* 2012;**51**:239–60.

Watanabe 2009

Watanabe N, Churchill R, Furukawa TA. Combined psychotherapy plus benzodiazepines for panic disorder. *Cochrane Database of Systematic Reviews* 2009, Issue 1. [DOI: 10.1002/14651858.CD005335.pub2]

White 2011

White IR. Multivariate random-effects meta-regression: updates to mvmeta. *The STATA Journal* 2011;**11**:255–70.

White 2012

White IR, Barrett JK, Jacksona D, Higgins JPT. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Research Synthesis Methods* 2012;**3**:111-25.

WHO 1992

World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Diagnostic Criteria for Research. Geneva: World Health Organization, 1992.

Wiborg 1996

Wiborg IM, Dahl AA. Does brief dynamic psychotherapy reduce the relapse rate of panic disorder?. *Archives of General Psychiatry* 1996;**53**:689-94.

Williams 1990

Williams SL. Guided mastery treatment of agoraphobia: beyond stimulus exposure. *Progress in Behavior Modification* 1990;**26**:89-121.

Wims 2010

Wims E, Titov N, Andrews G, Choi I. Clinician-assisted Internet-based treatment is effective for panic: a randomized controlled trial. *Australian and New Zealand Journal of Psychiatry* 2010;**44**(7):599-607.

Winston 2004

Winston A, Rosenthal RN, Pinsker H. Introduction to Supportive Psychotherapy. American Psychiatric Publishing, 2004.

Xiao 2011

Xiao Z, Li C, Wang J. Repetitive transcranial magnetic stimulation (rTMS) for panic disorder. *Cochrane Database of Systematic Reviews* 2011, Issue 4. [DOI: 10.1002/14651858.CD009083]

Zitrin 1978

Zitrin CM, Klein DF, Woerner MG. Behavior therapy, supportive psychotherapy, imipramine, and phobias. *Archives of General Psychiatry* 1978;**35**(3):307-16.

References to other published versions of this review

Pompoli 2014

Pompoli A, Furukawa T A, Imai H, Tajika A, Efthimiou O, Salanti G. Psychological therapies for panic disorder with or without agoraphobia in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 2.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Addis 2004

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** participants were eligible for the study if they met DSM-IV diagnostic criteria for panic disorder with or without agoraphobia or were subthreshold for a strict diagnosis of panic disorder but identified panic symptoms as their primary reason for seeking treatment (no subthreshold patient actually entered the study: *"seventy-three percent of participants met criteria for panic disorder with agoraphobia, and 27% met criteria for panic disorder without agoraphobia"*) **Exclusion criteria:** participants were excluded if they were seeking treatment for a problem other than panic or anxiety, had an untreated substance-use problem in the last 6 months, had a diagnosis of psychosis in the past 5 years, were currently judged to be at risk for suicide, or were concurrently involved in other individual psychotherapy. No exclusions were made on the basis of medication use for anxiety or other comorbid psychological or medical problems. **Characteristics of the sample:** • **Age:** mean age 39.9 years (SD 12.9, range 18 to 70) • **Percentage of agoraphobic patients:** 73% • **Percentage of patients on drug therapy:** 65% • **Percentage of patients with major depression:** 39%
Interventions	Participants (n = 80) were randomly assigned to either: **1) Panic control therapy** (classified as CBT, n = 38) • **Therapy format:** not stated • **Duration of each session:** not stated • **Mean number of sessions:** 12 to 15 • **Duration of intervention:** 12 to 15 weeks **2) Treatment as usual** (classified as SP, n = 42) • **Therapy format:** not stated • **Duration of each session:** not stated • **Mean number of sessions:** not stated

Addis 2004 *(Continued)*

	• **Duration of intervention:** not stated
Outcomes	**Time points for assessment:** baseline, 5.5 months, 8.5 months, 1 year, 2 years **Measures:** Panic Disorder Severity Scale (PDSS), Fear Questionnaire (FQ), Beck Depression Inventory (BDI–1), Outcome Questionnaire (OQ–45) The following outcomes were used for quantitative analyses: **ST-Remission:** PDSS below a cut-off score* at 5.5-month follow-up **ST-Response:** not measured (imputed from continuous scale) **ST-Dropouts:** patients who completed fewer than 6 sessions **Continuous scale:** PDSS at baseline and at 5.5 months **LT-Remission/Response:** PDSS below a cut-off score* at 1 year follow-up
Notes	** "Cut scores from published norms were obtained for the PDSS (Shear et al., 2001), the OQ–45 (Lambert et al., 1996), the FQ (Gillis, Haaga, & Ford, 1995), and the BDI–1 (Seggar, Lambert, & Hansen, 2002)"*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Low risk	All randomised patients (n = 80) were assessed at 5.5-month follow-up
Incomplete outcome data (attrition bias) Long-term	Low risk	All randomised patients (n = 80) were assessed at 1-year follow-up
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Low risk	*"Ten therapists agreed to participate. None of them identified their primary theoretical orientation as cognitive–behavioral; as a group they were approximately equally distributed between eclectic, family systems, psychodynamic, and humanistic in their self-described orientation."*
Treatment fidelity	Low risk	*"We rated therapist adherence for 67 of the 80 cases in the study. Data were missing for 11 cases in which the clients did not attend any treatment sessions and 2 cases in which the therapists had audiotaping difficulties. Cases with missing adherence data did not differ from the rest of the sample on any of the primary outcome measures at pre- or posttreatment.*

Addis 2004 *(Continued)*

PCT therapists scored higher than TAU therapists on all of the PCT interventions except for agoraphobic exposure, in which the frequency of use was low with no differences between the treatments."

Al Kubaisy 1992

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 60 years, agora-, social or specific phobic disorder on ICD-10 criteria for at least a year, mean 4-phobic-targets score of greater than 20 on a 0 to 32 scale, written informed consent (patients were told they would be in a research study about the best way to treat their kind of problem) **Exclusion criteria:** severe organic disease; failed exposure treatment in the last year; more than 2 units of alcohol a day from at least 3 weeks before entering the trial; on medication or on a stable dose of more than the daily equivalent of 5 mg of diazepam, 100 mg imipramine or 10 mg propranolol, taken only at night, for at least 4 months (by when it was unlikely to have any further effect, so this minority was retained to boost cell size) **Characteristic of the sample (agoraphobia sub-sample):** • **Age:** not specified • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 34) were randomly assigned to either: **1) Daily live self exposure homework + clinician accompanied live exposure** (classified as BT, n = 13) • **Therapy format:** not stated • **Duration of each session:** 150 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks **2) Daily live self exposure homework with six negotiation and monitoring sessions** (classified as BT, n = 11) • **Therapy format:** not stated • **Duration of each session:** 150 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks **3) Daily self relaxation homework with six negotiation and monitoring sessions** (classified as PT, n = 10) • **Therapy format:** not stated • **Duration of each session:** 150 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks
Outcomes	**Time points for assessment:** baseline, 8 weeks, 14 weeks, 26 weeks **Measures:** Fear Questionnaire (FQ), panic frequency, Beck Depression Inventory (BDI), Hamilton Depression (HAM-D), Global Improvement (CGI-I), Global Severity (CGI-S) The following outcomes were used for quantitative analyses:

Al Kubaisy 1992 *(Continued)*

	ST-Remission: not measured (imputed from continuous scale) **ST- Response:** not measured for agoraphobia sub-sample (imputed from continuous scale) **ST-Dropouts:** refusers and dropouts before week 8 **Continuous scale:** although measured (CGI) detailed data are not reported **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	ST-Remission not measured: imputed from CGI-S, which was rated by *"an assessor (psychiatrists, psychologists and nurse therapists) kept blind to the treatment condition."*
Incomplete outcome data (attrition bias) Short-term	High risk	29% of randomised patients (agoraphobia sub-sample) did not receive/complete the assigned intervention. Dropouts imbalanced in number across the 3 arms. Refusers and dropouts data not reported.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"Both Ee and e patients had the rationale of exposure explained at the 1st session, and were asked to read the self-help chapter from Living With Fear (Marks IM, 1980) and to follow its guidelines".* Marks IM is among the study authors.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided
Other bias	Unclear risk	Modification of the original sample with replacements. Number and randomisation of replacements not specified.

Arch 2012

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM–IV diagnosis of one or more anxiety disorders, including panic disorder with or without agoraphobia (PD/A), social anxiety disorder (SAD), specific phobia (SP), obsessive-compulsive disorder (OCD), or generalised anxiety disorder (GAD) **Exclusion criteria:** none **Characteristic of the sample (agoraphobia sub-sample):** • **Age:** not specified • **Percentage of agoraphobic patients:** not specified

Arch 2012 *(Continued)*

- **Percentage of patients on drug therapy:** not specified
- **Percentage of patients with major depression:** not specified

Interventions	Participants (n = unclear) were randomly assigned to either: **1) Acceptance and commitment therapy** (classified as 3W, n = unclear) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) Cognitive behaviour therapy** (classified as CBT, n = unclear) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** pre-treatment, post-treatment, 6 months follow-up, 12 months follow-up **Measures:** Anxiety Disorders Interview Schedule–IV (ADIS–IV), Anxiety Sensitivity Index (ASI), Penn State Worry Questionnaire (PSWQ), Fear Questionnaire (FQ), Main Target Phobia Scale (a single-item avoidance rating for each participant's "main phobia"), Quality of Life Inventory (QOLI), Acceptance and Action Questionnaire–16 (AAQ) The following outcomes were used for quantitative analyses: **ST-Remission:** although reported (Clinical Severity Rating on ADIS-IV lower than 4), ST-Remission could not be calculated following an ITT principle (number of PD/A patients randomised to each arm not specified) **ST- Response:** not measured **ST-Dropouts:** not measured **Continuous scale:** ASI at pre- and post-treatment **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"Randomization sequences were produced by http://www.randomizer.org"*
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"One hundred and forty-three participants [...] were randomized to ACT (n=65) or CBT (n=78). All participants who began treatment (n=128) were included in the intent-to-treat (ITT) sample (n 57=ACT, n=71 CBT)".* Data for randomised patients who did not begin treatment unavailable.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable

Arch 2012 *(Continued)*

Researcher allegiance	Unclear risk	*"CBT for anxiety disorders followed a protocol authored by Craske"; "ACT for anxiety disorders followed a manual authored by Eifert and Forsyth".* Both Craske MG and Eifert GH are among the study authors: although possible, the direction of a researcher allegiance bias would be unclear.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All sessions were videotaped for supervision purposes with a hidden video camera; sessions were also audiotaped for therapy adherence purposes with a discrete digital recorder. Videos were generally played in supervision sessions or watched beforehand by supervisors."*

Arntz 2002

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 17 and 70 years, primary DSM-III-R diagnosis of panic disorder with no or mild agoraphobic avoidance, panic disorder for at least 3 months, at least one panic attack during the previous 4 weeks, asking for treatment of panic disorder. No use of serotonergic antidepressants or benzodiazepines (for at least 4 weeks; patients using this medication were, if they agreed, taken off medication), or if unwilling to stop medication, keeping this medication at a constant level during treatment or stopping it during treatment. **Exclusion criteria:** depressive disorder preceding the current episode of panic disorder or requiring immediate treatment; behaviour therapy received for panic disorder; evidence of organic mental disorders, mental retardation, psychotic disorders, alcohol or drug dependence, cardiovascular disease, asthma, epilepsy; medical contraindication for exposure, behavioural experiments or hyperventilation. **Characteristic of the sample:** • **Age:** 34.8 years (range 20 to 65) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** 27.5% • **Percentage of patients with major depression:** not specified (mood disorder 33.3%)
Interventions	Participants (n = 69) were randomly assigned to either: **1) Cognitive therapy** (classified as CT, n = unclear) • **Therapy format:** individual/group • **Duration of each session:** 60 minutes (individual sessions)/105 minutes (group sessions) • **Mean number of sessions:** 16 • **Duration of intervention:** 14 weeks (plus 2 sessions after 1 and 6 months) **2) Interoceptive exposure** (classified as BT, n = unclear) • **Therapy format:** individual/group • **Duration of each session:** 60 minutes (individual sessions)/105 minutes (group sessions) • **Mean number of sessions:** 16 • **Duration of intervention:** 14 weeks (plus 2 sessions after 1 and 6 months)
Outcomes	**Time points for assessment:** baseline, post-treatment (16 weeks), 1-month follow-up, 6 months follow-up **Measures:** panic diary, Fear of Fear Questionnaire, Fear Questionnaire (FQ), State–Trait Anxiety Inventory (STAI), Symptom Check List (SCL-90)

Arntz 2002 *(Continued)*

	The following outcomes were used for quantitative analyses: **ST-Remission:** although measured (panic-free at post-treatment), data cannot be used to calculate remission following an ITT principle (n randomised for each arm is unclear) **ST- Response:** not measured **ST-Dropouts:** detailed data are not reported **Continuous scale:** although measured, detailed data are not reported **LT-Remission/Response:** although measured (panic-free at 6 months follow-up) detailed data are not reported
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Selective reporting (reporting bias)	High risk	Study protocol unavailable. Pre-planned measures are not reported with sufficient details.
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"There were weekly supervisions, during which each case was presented in detail and adherence to the protocol was checked. Care was taken to exclude cognitive techniques from the IE treatment, and exposure techniques from the CT treatment."*

Barlow 1989

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM III-R diagnosis of panic disorder with mild or no agoraphobic avoidance. The interviewers rated the severity of the disturbance on a 0 to 8-point scale (reflecting co-jointly distress and disability from the disorder), and only clients whose severity rating was at least 4 were included in the study. Finally, only subjects who reported the presence of at least 1 panic attack in a 2-week period prior to assessment were included. Subjects on medications or receiving alternative psychotherapies for the requisite time, and who met suitability criteria, were included under the agreement that medication regime and psychotherapy contact were maintained at constant levels throughout. **Exclusion criteria:** aged below 18 or above 65 years; current alcohol or drug dependency/abuse; primary diagnosis of major depression, and any signs of psychosis or organic brain syndrome. In addition, clients concurrently involved in other psychotherapy programs were assessed for suitability only if the alternative therapy was not focused on anxiety management, and they had been in therapy for at least 6 months. Finally, subjects were excluded if they had begun benzodiazepines within the past 3 months or MAO inhibitors or tricyclic antidepressants within the past 6 months.

Barlow 1989 *(Continued)*

Characteristic of the sample:

- **Age:** not specified for the ITT sample (completers sub-sample mean age 31.7 years, SD 8.3)
- **Percentage of agoraphobic patients:** not specified
- **Percentage of patients on drug therapy:** not specified for the ITT sample (25% in completers sub-sample)
- **Percentage of patients with major depression:** not specified for the ITT sample (5% in completers sub-sample)

Interventions	Participants (n = 71) were randomly assigned to either: **1) Wait list** (classified as WL, n = 16) • **Mean number of sessions:** 0 • **Duration of intervention:** 15 weeks **2) Applied progressive muscle relaxation** (classified as PT, n = 15) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks **3) Exposure and cognitive restructuring** (classified as CBT, n = 16) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks **4) Relaxation combined exposure and cognitive restructuring** (classified as CBT, n = 24) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks
Outcomes	**Time points for assessment:** assessments were conducted at pre-treatment and post-treatment. Active treatment group subjects were also assessed 3 months, 6 months, 12 months and 24 months after treatment completion. **Measures:** Anxiety Disorder Interview Schedule-Revised (ADIS-R), Trait Scale of the State-Trait Anxiety Inventory (STAI-T), Cognitive-Somatic Anxiety Questionnaire, Fear Questionnaire (FQ), Beck Depression Inventory (BDI), Psychosomatic Rating Scale, Subjective Symptom Scale, self monitoring records (regarding daily fluctuations in anxiety and depression and occurrence of panic attacks), composite criteria for treatment responder (20% improvement in at least 3 of the following 4 measures: (1) clinical rating of severity (at least 2 points); (2) client's self rating from the Fear Questionnaire (at least 2 points); (3) number of panic attacks per week, and (4) Subjective Symptom Scale total score (at least 8 points) and End-State Functioning (applied only to treatment responders. At least three of the following five criteria had to be obtained for high end-state status: (1) score of 2 or less on the clinician's rating of severity; (2) score of 2 or less for the client's self rating; (3) 0 panic attack per week; (4) score of 2 or less for the mean anxiety rating, and (5) score of 10 or less for the Subjective Symptom Scale total score). The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* at post-treatment **ST- Response:** treatment responder as defined by composite criteria **ST-Dropouts:** subjects who did not complete assigned treatment

Barlow 1989 *(Continued)*

	Continuous scale: Fear Questionnaire (FQ) **LT-Remission/Response:** reported data were not included in the analyses because of high level of dropouts (see Secondary outcomes): *"data were available for 23 subjects at the 6-month follow-up period (R n=9; E&C n=8; COMB n=6)"*
Notes	* *"At least three of the following five criteria had to be obtained for high end state status: (1) score of 2 or less on the clinician's rating of severity; (2) score of 2 or less for the client's selfrating; (3) zero panic attack per week; (4) score of 2 or less for the mean anxiety rating, and (5) score of 10 or less for the Subjective Symptom Scale total score. End state functioning was determined if data from only three different measures were present but all three reflected positive or negative responding. End state status could not be determined if more than two of the five measures were missing."*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	At least 3 out of 5 criteria had to be obtained for high end-state status. Although assessor was a *"blind, independent rater"*, only one of those 5 criteria was assessor-rated, so it was possible for a patient to fall into the end-state category on the basis of self rated measures only.
Incomplete outcome data (attrition bias) Short-term	High risk	"*The percentage of dropouts for each condition were 6%, 33%, 6%, and 17%. For the study completers, data were missing at post-test for several variables due to non-compliance. The number of missing data points ranged from 1 to 4 variables within each group. Missing data were not replaced by averages.*"
Incomplete outcome data (attrition bias) Long-term	High risk	***"****data were available for 23 subjects at the 6-month follow-up period (R n=9; E&C n=8; COMB n=6)."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Barlow DH and Craske MG are authors of a CBT manual (see Barlow 2000b)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"Treatment delivery was examined by means of ratings of the content of therapy sessions from periodic spot checks of audiotapes (all therapy sessions were audiotaped to avoid the possibility of response bias in the therapists verbal behavior during spot checking). Thirty-five tapes were randomly selected, with the stipulation that each therapist and each treatment phase of each treatment condition were represented in the sample. Two randomly selected five minute segments (excluding the first and last five minutes of the session and including at least three minutes of therapist talk) were rated from each tape. In all cases, raters identified correctly the treatment condition represented by the sample. Judgments of the treatment phase from which the sample came were correct in 31 of the 35 cases; two misjudgements were from the E & C condition and two from the R condition. There were only two instances of inappropriate material; both of which referred to nontargeted problem areas and not to inappropriate treatment technique."*

Beck 1987

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** not specified, probably DSM-III diagnosis of panic disorder (as in another previous study by Ottaviani and Beck described in the book) **Exclusion criteria:** not specified **Characteristic of the sample:** • **Age:** not specified • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 29) were randomly assigned to either: **1) Cognitive therapy** (classified as CBT, n = 13) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) Brief supportive therapy** (classified as APP, n = 16) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks
Outcomes	**Time points for assessment:** baseline, 4 weeks, 8 weeks, 12 weeks (only for cognitive therapy group) **Measures:** panic frequency The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (not imputed because of skewed distribution of available continuous scale) **ST- Response:** not measured (not imputed because of skewed distribution of available continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** although measured (panic frequency), detailed data are not reported **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided

Beck 1987 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Low risk	*"There were no dropouts in either group".*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Beck AT is involved in conceptualisation of cognitive therapy (see Description of the intervention)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Beck 1992

Methods	**Study design:** randomised controlled trial, cross-over design
Participants	**Inclusion criteria:** age between 18 and 65 years, DSM-III diagnosis of panic disorder or agoraphobia with panic disorder **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** not specified • **Percentage of agoraphobic patients:** 18% • **Percentage of patients on drug therapy:** 52% • **Percentage of patients with major depression:** not specified (35% according to imputation from BDI)
Interventions	Participants (n = 33) were randomly assigned to either: **1) Focused cognitive therapy** (classified as CT, n = 17) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) Brief supportive psychotherapy** (classified as WL*, n = 16) • **Therapy format:** not specified • **Duration of each session:** 30 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks
Outcomes	**Time points for assessment:** baseline, 4 weeks, 8 weeks, 12 weeks (only for focused cognitive therapy group) **Measures:** panic frequency, Mobility Inventory for Agoraphobia (MI), Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), Specific Fear Inventory, end-state functioning (only for focused cognitive therapy group)

Beck 1992 *(Continued)*

	The following outcomes were used for quantitative analyses: **ST-Remission:** absence of panic attacks (clinician rating, before cross-over) **ST- Response:** not measured (not imputed) **ST-Dropouts:** non-completers (before cross-over) **Continuous scale:** not extracted (number of assessed patients unclear: *"Ns varied across analyses from 14 to 17 patients in the cognitive therapy group and 15 to 16 patients in the brief supportive psychotherapy group"*) **LT-Remission/Response:** not measured
Notes	*Brief supportive psychotherapy arm was classified as WL (wait list) because: 1) patients received *"8 weeks of supportive contact"*, apparently different from a proper supportive therapy (which therefore cannot be classified as an active treatment, but rather as a comparator intervention); 2) although this study is presented as having a cross-over design, only patients in supportive psychotherapy group (all of them) actually did cross-over.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	*"At each assessment interval, independent clinical raters reviewed patients' daily logs of panic frequency to determine whether the recorded panic attacks actually met the DSM-III criteria for panic"*. It is unclear whether raters were blind to patients' allocation.
Incomplete outcome data (attrition bias) Short-term	High risk	*"Because of missing values for some variables, Ns varied across analyses from 14 to 17 patients in the cognitive therapy group and 15 to 16 patients in the brief supportive psychotherapy group"*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Beck AT is involved in conceptualisation of cognitive therapy (see Description of the intervention)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Beck 1994

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age 18 to 65 years, DSM-III primary diagnosis of panic disorder. Patients who reported use of psychotropic medication were withdrawn from these regimes, with at least a 2-week drug clearance before taking the ADIS-R.

Beck 1994 *(Continued)*

	Exclusion criteria: severe agoraphobia, primary diagnosis of an alternate Axis I diagnosis, current involvement in psychotherapy, alcohol or substance abuse within the previous 6 months, psychotic symptoms, evidence of organic impairment. **Characteristic of the sample:** • **Age:** not specified for the randomised sample (n = 64) but only for the initially selected sample (n = 70, mean age 37.5 years, SD 9.7) • **Percentage of agoraphobic patients:** not specified for the randomised sample (n = 64) but only for the initially selected sample (n = 70, 87% being moderately or mildly agoraphobic) • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified for the randomised sample (n = 64) but only for the initially selected sample. Depression percentage among randomised sample was 23% according to imputation based on HAM-D-17 score)
Interventions	Participants (n = 64) were randomly assigned to either: **1) Cognitive therapy** (classified as CT, n = 22) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks **2) Relaxation training** (classified as PT, n = 20) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks **3) Minimal contact control** (classified as NT, n = 22) • **Therapy format:** weekly telephone contact • **Duration of each session:** not specified • **Mean number of sessions:** 0 • **Duration of intervention:** 10 weeks
Outcomes	**Time points for assessment:** baseline, 5 weeks, 10 weeks. Subjects in the 2 intervention groups were asked to return for 1-, 3- and 6-months follow-up visits. **Measures:** panic frequency, Anxiety Disorders Interview Schedule-Revised (ADIS-R), Anxiety Sensitivity Index (ASI), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), State-Trait Anxiety Inventory (STAI), Fear Questionnaire (FQ), Hamilton Anxiety and Depression Scales (HAM-A, HAM-D), composite index of treatment response (based on 4 variables: global PD severity, number of panic attacks in the previous month, average ACQ and BSQ score, FQ-Ag score). The following outcomes were used for quantitative analyses: **ST-Remission:** panic-free at post-test (10 weeks) **ST- Response:** at least mild improvement on composite index of treatment response* **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** panic-free at 6 months follow-up
Notes	* *"A composite index of treatment response was derived, using guidelines established by Himadi, Boice, and Barlow (1986) and Barlow et al. (1989). This measure quantified treatment response based on four*

Beck 1994 *(Continued)*

variables: global PD severity, number of panic attacks in the previous month, an average of ACQ and BSQ scores, and FQ-Ag score".

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"Six subjects dropped out, with 5 (23%) from the CT condition, 1 (5%) from the RT condition, and none (0%) from the MCC condition".*
Incomplete outcome data (attrition bias) Long-term	High risk	*"Of the 17 subjects who completed CT, 16 (94%) were assessed at all three follow-up points, with 1 subject not assessed at 3 and 6 months. All 19 RT subjects were evaluated at each follow-up assessment".*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapies under study (note that first author is Beck JG, not Beck AT).
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"We provided weekly clinical supervision for each session to ensure treatment competence. All sessions were videotaped, with 24% (n=34) selected randomly for treatment integrity monitoring"*

Beutel 2013

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age 18 to 60 years, primary panic disorder with or without agoraphobia according to DSM-IV criteria, fluency in the German language, living in the proximity of Mainz. Psychotropic medication, if present, had to be held constant. **Exclusion criteria:** psychosis, bipolar disorder, borderline or antisocial personality disorder, active substance abuse, severe medical or neurological disease precluding exposure therapy and ongoing psychotherapy. **Characteristic of the sample:** • **Age:** 36.22 years (SD 10.8) • **Percentage of agoraphobic patients:** 74.1% • **Percentage of patients on drug therapy:** 22.2% • **Percentage of patients with major depression:** 22.2%

Beutel 2013 *(Continued)*

Interventions	Participants (n = 54) were randomly assigned to either: **1) Panic focused psychodynamic therapy** (classified as PD, n = 36) • **Therapy format:** not specified • **Duration of each session:** 50 minutes • **Mean number of sessions:** 24 • **Duration of intervention:** 12 weeks **2) Cognitive behaviour therapy** (classified as CBT, n = 18) • **Therapy format:** not specified • **Duration of each session:** 50 minutes • **Mean number of sessions:** 24 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, treatment termination, 6 months follow-up **Measures:** Structured Clinical Interview for DSM-IV (SCID-I and II), Panic Disorder Severity Scale (PDSS), Clinical Global Impression scale (CGI), Hamilton Depression Rating Scale (HAM-D), Symptom Checklist (SCL-90R), Beck Depression Inventory (BDI), Levels of Emotional Awareness Scale (LEAS) The following outcomes were used for quantitative analyses: **ST-Remission:** PDSS score < 5 in PD or < 7 in PDA at termination **ST- Response:** at least 40% reduction of PDSS **ST-Dropouts:** non-completers **Continuous scale:** Panic Disorder Severity Scale (PDSS) **LT-Remission/Response:** PDSS score < 5 in PD or < 7 in PDA at 6 months follow-up
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"subjects were allocated by an a priori computer-generated list in a 2: 1 randomization ratio either to Panic Focused Psychodynamic Psychotherapy or to manualized cognitive behavioral therapy plus exposure."*
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	"*Independent, experienced evaluators, blinded to subject condition and therapist orientation, assessed subjects at baseline, at treatment termination and at 6 months follow-up.*"
Incomplete outcome data (attrition bias) Short-term	High risk	Patients analysed at post-treatment: PFPP n = 28 CBT n = 14
Incomplete outcome data (attrition bias) Long-term	High risk	Patients analysed at follow-up: PFPP n = 25

Beutel 2013 *(Continued)*

		CBT n = 13
Selective reporting (reporting bias)	High risk	Study protocol available (registered retrospectively). Reported primary outcome is one among other primary outcomes cited in the protocol: *"Principal outcome criterion is the reduction of panic-related symptoms at the follow-up 6-months after treatment. Panic-related symptoms are measured with standardized questionnaires and interviews, e. g. the AKV-MI/BSQ/ACQ questionnaires, the Hamilton Anxiety Scale, HAMA, and the Panic Disorder Severity scale, PDSS."*
Researcher allegiance	High risk	Milrod B is co-author of PFPP manual (see Milrod 1997)
Therapist allegiance	Unclear risk	Insufficient information provided
Treatment fidelity	Low risk	*"All treatments were videotaped as a basis for supervision and for later independent assessment of treatment adherence"*

Botella 2004

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** 18 years of age or older, met DSM-IV (American Psychiatric Association, 2000) criteria for the diagnosis of PDA as principal diagnosis and, in the case of taking medication for PDA, did not increase or modify the kind of medication during the study. **Exclusion criteria:** psychosis, severe organic illness or substance abuse or dependence **Characteristic of the sample:** • **Age:** mean 34.7 years (SD 12.31) • **Percentage of agoraphobic patients:** 82.9% • **Percentage of patients on drug therapy:** 66.6% • **Percentage of patients with major depression:** not reported (29.7% according to imputation from BDI)
Interventions	Participants (n = 37) were randomly assigned to either: **1) In vivo exposure** (classified as CBT, n = 12) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 9 • **Duration of intervention:** 9 weeks **2) Virtual reality exposure** (classified as CBT, n = 12) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 9 • **Duration of intervention:** 9 weeks **3) Wait list** (classified as WL, n = 13) • **Mean number of sessions:** 0 • **Duration of intervention:** 9 weeks
Outcomes	**Time points for assessment:** pre-treatment, post treatment, 12 months follow-up

Botella 2004 *(Continued)*

	Measures: Anxiety Diagnostic Interview Schedule IV (ADIS-IV), Fear and Avoidance Scales, panic attack record, Panic Disorder Severity Scale (PDSS), Anxiety Sensitivity Index (ASI), Agoraphobia Subscale of the Fear Questionnaire (FQ-Ag), Beck Depression Inventory (BDI), Maladjustment Scale (MS), Clinician Global Impression (CGI) The following outcomes were used for quantitative analyses: **ST-Remission:** ASI score < 27 **ST- Response:** panic-free OR a 50% reduction in panic frequency **ST-Dropouts:** non-completers **Continuous scale:** Panic Disorder Severity Scale (PDSS) **LT-Remission/Response:** not entered in the analyses (reported LT data refer to the 2 treatment arms, both classified as CBT in this review: comparison not feasible)
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random numbers table was used (personal communication)
Allocation concealment (selection bias)	Low risk	Randomisation was performed by an experimenter who did not participate in the study (personal communication)
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Assessors were blind to the conditions (personal communication). However, ASI (used to determine ST-Remission) is a self administered scale.
Incomplete outcome data (attrition bias) Short-term	Low risk	*"All participants in the treatment conditions were assessed at post treatment 1 week after the treatment completion"*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Authors involved in the developing of virtual reality exposure treatment for panic disorder (*"This finding has encouraged us to design a VRE treatment for PDA. Our VR programme for PDA includes several VR scenarios"*)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"The therapists were well trained in CBT programmes for PDA. Treatment adherence across the therapists was ensured by a specific training in the treatment programmes. Also, the complete team held weekly meetings to supervise the ongoing treatment of all patients"*

Brown 1997

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 65 years, DSM-III-R diagnosis of panic disorder with or without agoraphobia, at least 1 panic attack in the month preceding the intake evaluation

Brown 1997 *(Continued)*

	Exclusion criteria: actively psychotic, immediate suicidal or homicidal risk, current abuse of any substance, brain-damage, in the manic phase of a bipolar disorder without medication **Characteristic of the sample:** • **Age:** not specified for ITT sample (among the 40 completers, mean age was 33 years, SD 9.8, range 19 to 56) • **Percentage of agoraphobic patients:** not specified for ITT sample (75% among the 40 completers) • **Percentage of patients on drug therapy:** not specified for ITT sample (52.5% among the 40 completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 48) were randomly assigned to either: **1) Focused cognitive therapy** (classified as CBT, n = unclear) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 14 • **Duration of intervention:** 14 weeks **2) Standard cognitive therapy** (classified as CT, n = unclear) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 14 • **Duration of intervention:** 14 weeks
Outcomes	**Time points for assessment:** baseline, termination, 6 months follow-up, 12 months follow-up **Measures:** panic frequency, Hamilton Anxiety Rating Scale-Revised (HARS-R), Hamilton Rating Scale for Depression-Revised (HRSD-R), Beck Anxiety Inventory (BAI), Beck Depression Inventory (BDI), Agoraphobic Cognition Questionnaire (ACQ), Panic Belief Questionnaire (PBQ) The following outcomes were used for quantitative analyses: **ST-Remission:** although reported (panic-free at termination), ST-Remission could not be calculated following an ITT principle (number of patients randomised to each arm not specified) **ST- Response:** not measured (not imputed) **ST-Dropouts:** not measured **Continuous scale:** panic frequency **LT-Remission/Response:** although reported (panic-free at 12 months follow-up), ST-Remission could not be calculated following an ITT principle (number of patients randomised to each arm not specified)
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided

Brown 1997 *(Continued)*

Incomplete outcome data (attrition bias) Short-term	High risk	*"Forty-eight subjects initially agreed to participate in the study and 40 patients successfully completed the study through the l-year follow-up period. Of eight patients who dropped out of the study, one patient was hospitalized for medical reasons after receiving five SCT sessions, one patient failed to complete a significant portion of the outcome measures at several assessment points, one patient decided to pursue pharmacotherapy exclusively for panic disorder, three patients decided to pursue alternative psychotherapy interventions, and two patients dropped out of the study for unknown reasons."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"Focused cognitive therapy for panic disorder was developed from a theoretical model of panic disorder (Beck, Emery, & Greenberg, 1985; Clark, 1986)."* Beck AT is among the study authors.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	High risk	"*Results of tape ratings indicated that all therapists addressed catastrophic interpretations according to protocol for the 21 patients in the FCT group. However, protocol violations were noted for 8 of the 19 patients who received SCT.*"

Burke 1997

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** females, primary DSM-III diagnosis of agoraphobia. Patients were instructed not to change their medication during the trial. **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** in exposure group, mean age was 40 years, SD 8.9; in CBT group, mean age was 40.1, SD 11.08 • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified for ITT sample (53.8% among the 26 completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 39) were randomly assigned to either: **1) Exposure** (classified as BT, n = 20) • **Therapy format:** individual • **Duration of each session:** 150 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks **2) Cognitive behaviour therapy** (classified as CBT, n = 19) • **Therapy format:** individual • **Duration of each session:** 180 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 6 months follow-up

Burke 1997 *(Continued)*

	Measures: Fear Questionnaire (FQ), Agoraphobia Questionnaire, Spielberger Trait Anxiety Inventory (STAI), Beck Depression Inventory (BDI), behavioural test, Agoraphobic Cognitions Questionnaire (ACQ), Anxiety Scale of the Cognitions Checklist (CCLAS), Probability Questionnaire (PQ), Evaluation Questionnaire (EQ) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	ACQ (used to determine ST-Remission) is a self administered scale
Incomplete outcome data (attrition bias) Short-term	High risk	*"Thirteen participants dropped out of treatment."* Reported data refer to treatment completers.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapies under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"CBT session and a tape of an Exp session from each therapist to send to an independent assessor not involved in the trial. The assessor had been trained to teach CT at the Center for Cognitive Therapy, Philadelphia and she teaches a specialist training course in CT in the UK.' A table of random numbers was used to select which of each therapist's CBT and Exp tapes were sent to the assessor. The total sample of tapes came to 18 (10 CBT and eight Exp).*"

Carter 2003

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** African American population, DSM-IV diagnosis of panic disorder with agoraphobia

Carter 2003 *(Continued)*

	Exclusion criteria: any psychotic disorder, current substance abuse or dependence, significant suicidal ideation/gestures, any comorbid condition receiving a clinical severity rating equal to or greater than that assigned the panic disorder diagnosis **Characteristic of the sample:** • **Age:** not specified for ITT sample (among the 25 completers, mean age was 42.36 years, SD 6.7, for treatment group; mean age was 4.55 years, SD 5.5, for wait list) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified for ITT sample (among the 25 completers, 84.5% had comorbid depression)
Interventions	Participants (n = 32) were randomly assigned to either: **1) Panic control treatment** (classified as CBT, n = 17) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 11 • **Duration of intervention:** 11 weeks **2) Wait-list control**(classified as WL, n = 15) • **Mean number of sessions:** 0 • **Duration of intervention:** 11
Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** Structured Clinical Interview for DSM-IV Axis I Disorders (ADIS-IV), Anxiety Sensitivity Index (ASI), State-Trait Anxiety Inventory (STAI), Beck Depression Inventory (BDI), Beck Depression Inventory (BDI), The Hyperventilation Questionnaire - Cognitive Subscale (HQC), The Mobility Inventory (MI), African American Acculturation Scale - Short Form (AAAS), Attitude Toward Treatment Questionnaire (ATQ) The following outcomes were used for quantitative analyses: **ST-Remission:** recovery (based on ASI)* **ST- Response:** improvement + recovery (based on ASI)* **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	* *"To assess clinically significant change, the method described by Jacobson and Truax (1991) was employed. For each dependent variable (except the HQC) for which there were established cutoff scores and test–retest reliability statistics, the level of functioning following therapy for each patient was examined and judged whether it was closer to the mean of a functional population than it was to the dysfunctional population. As suggested by Jacobson and Truax (1991), a reliable change (RC) index was computed for each group to assess whether fluctuations were likely the result of imprecise measurement. Each patient was categorized as recovered (score closer to the mean of the functional than dysfunctional group and RC greater than 1.96), improved but not recovered (score closer to the mean of the functional group, but the change noted did not exceed the RC cutoff of 1.96), or unimproved."*

Risk of bias

Bias	Authors' judgement	Support for judgement

Carter 2003 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	ASI (used to determine ST-Remission) is a self administered scale
Incomplete outcome data (attrition bias) Short-term	High risk	*"Three of the 17 treatment patients were classified as non-completers. Of the 15 assigned to the wait-list condition, 4 did not return for the second evaluation (26.6% attrition). We report the data from the remaining 25 patients who completed either treatment or the wait-list assessment."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	*"The lead therapist for all groups was a licensed clinical psychologist who is an African American male with 15 years experience with cognitive behavioral therapy for anxiety disorders."*
Treatment fidelity	Unclear risk	No information provided

Clark 1994

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age 18 to 60 years, DSM-III-R diagnosis of panic disorder with no, mild or moderate agoraphobic avoidance, current episode duration at least 6 months (this criterion was intended to minimise spontaneous remission, at least 3 panic attacks in the last 3 weeks, consider panic their main problem, willing to accept random allocation. **Exclusion criteria:** depressive disorder severe enough to require immediate psychiatric treatment; cognitive therapy, applied relaxation or imipramine in the current episode; evidence of organic mental disorder, schizophrenia, alcohol or drug dependence, cardiovascular disease, asthma, epilepsy; pregnancy or intention to become pregnant. Concurrent Axis II personality disorder was not a reason for excluding patients unless personality disorder was clearly the primary problem. **Characteristic of the sample:** • **Age:** not specified for ITT sample (among the 64 completers, mean age was 34.6 years, SD 9.2) • **Percentage of agoraphobic patients:** not specified for ITT sample (among the 64 completers, 81% had mild or moderate agoraphobic avoidance) • **Percentage of patients on drug therapy:** unclear • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 72) were randomly assigned to either: **1) Cognitive therapy + in vivo exposure** (classified as CBT, n = 17, n = 21 after re-randomisation of WL patients) • **Therapy format:** not specified • **Duration of each session:** 60 minutes

Clark 1994 *(Continued)*

- **Mean number of sessions:** 15
- **Duration of intervention:** 24 weeks

2) Applied relaxation + in vivo exposure (classified as BT, n = 17, n = 21 after re-randomisation of WL patients)

- **Therapy format:** not specified
- **Duration of each session:** 60 minutes
- **Mean number of sessions:** 15
- **Duration of intervention:** 24 weeks

3) Imipramine + in vivo exposure (not included in this review, n = 22, n = 26 after re-randomisation of WL patients)

- **Therapy format:** not specified
- **Duration of each session:** 25 minutes
- **Mean number of sessions:** 15
- **Duration of intervention:** 24 weeks

4) Wait list (classified as WL, n = 16; after waiting period, 12 patients were re-randomised to 1 of the 3 active treatments)

- **Mean number of sessions:** 0
- **Duration of intervention:** 12 weeks

Outcomes	**Time points for assessment:** baseline, 3 months follow-up, 6 months follow-up, 15 months follow-up **Measures:** panic frequency, panic-related distress/disability, Beck Anxiety Inventory (BAI), Hamilton Anxiety Rating Scale (HARS), Fear Questionnaire (FQ), Bodily Sensations Questionnaire (BSQ), Body Sensations Interpretation Questionnaire (BSIQ), Agoraphobic Cognitions Questionnaire (ACQ), Beck Depression Inventory (BDI) The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* at 3 months (original sample + re-randomised WL patients) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers (< 3 sessions) at 3 months (original sample) **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ), measured on original sample + re-randomised WL patients **LT-Remission/Response:** high end-state functioning* at 15 months (original sample + re-randomised WL patients)
Notes	*High end-state function was defined as panic-free and an assessor panic-related distress/disability rating equal or below 2 ('slight').

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias)	Low risk	*"Assessments, which included ratings completed by an assessor who was blind to treatment allocation, were at pre-treatment/waiting-list, 3, 6, and 15 months."*

Clark 1994 *(Continued)*
ST-Remission

Incomplete outcome data (attrition bias) Short-term	Low risk	*"Of 72 patients meeting acceptance criteria,3 dropped out (1per treatment). Five who agreed initially to random allocation refused to take imipramine when allocated to that condition. Drop-outs and refusers after randomisation were replaced and not included in the data analysis.To be classified as a drop-out, patients had to start treatment but attend no more than two sessions. Patients who attended at least three sessions were considered completers and included in all analyses."* For this review we are not considering the imipramine arm, therefore dropouts (CBT n = 1; BT n = 1; WL n = 0) were low in number and evenly distributed. Therefore, the proportion of missing outcomes is not enough to have a clinically relevant impact on the intervention effect estimate.
Incomplete outcome data (attrition bias) Long-term	Low risk	(See above)
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"Cognitive therapy (CT) was based on the cognitive theory of panic. Several cognitive and behavioural techniques (see Clark, 1989; Salkovskis & Clark, 1991) were used to help patients identify and change misinterpretations of bodily sensations."* Both Clark DM and Salkovskis PM are among the study authors.
Therapist allegiance	Unclear risk	Insufficient information provided
Treatment fidelity	Low risk	*"To check therapists' adherence to the treatment protocol ten audiotapes per treatment (each from a different patient) were randomly selected and rated for the presence/absence of features which should be unique to that treatment and for time spent on procedures which should be common to all treatments. There were no protocol violations and the treatments did not differ in times spent on the common procedures."*

Clark 1999

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age 18 to 60 years, DSM-III-R diagnosis of panic disorder with no, mild or moderate agoraphobic avoidance, current episode duration at least 6 months (this criterion was intended to minimise spontaneous remission, at least 3 panic attacks in the last 3 weeks, consider panic their main problem, willing to accept random allocation, no use of medication (or, if taking psychotropic medication, on a stable dose for at least 3 months with an agreement not to change dosage), record of at least one panic attack while keeping a daily panic diary during a post-interview 2-week baseline period. **Exclusion criteria:** depressive disorder severe enough to require immediate psychiatric treatment; previous treatment with cognitive therapy or exposure therapy for panic disorder; evidence of organic mental disorder, schizophrenia, alcohol or drug dependence, cardiovascular disease, asthma, epilepsy; pregnancy or intention to become pregnant. **Characteristic of the sample:** • **Age:** mean age 34 years (SD 11.1) • **Percentage of agoraphobic patients:** 85% • **Percentage of patients on drug therapy:** 12%

Clark 1999 *(Continued)*

	• **Percentage of patients with major depression:** not specified (30.2% imputed from BDI)
Interventions	Participants (n = 43) were randomly assigned to either: **1) Full cognitive therapy** (classified as CBT, n = 15) • **Therapy format:** not specified • **Duration of each session:** 66 minutes (average) • **Mean number of sessions:** 15 • **Duration of intervention:** 24 weeks **2) Brief cognitive therapy** (classified as CBT, n = 14) • **Therapy format:** not specified • **Duration of each session:** 71.25 (average) • **Mean number of sessions:** 8 • **Duration of intervention:** 24 weeks **3) Wait list** (classified as WL, n = 14) • **Mean number of sessions:** 0 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment/wait list, 3 months post-treatment follow-up, 12 months post-treatment follow-up **Measures:** panic-anxiety composite measure, panic frequency, panic-related distress/disability, Beck Anxiety Inventory (BAI), Hamilton Anxiety Rating Scale (HARS), Fear Questionnaire (FQ), Body Sensations Interpretation Questionnaire (BSIQ), Agoraphobic Cognitions Questionnaire (ACQ), Beck Depression Inventory (BDI) The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ) **LT-Remission/Response:** although measured (high end-state functioning*), data could not be extracted because the comparison was between the 2 active treatment arms, both classified as CBT (comparison not feasible).
Notes	* *"We defined high end-state functioning as panic free and as an assessor-scored panic-related distress—disability rating of 2 or less ('slight')"*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	*"Assessments, which included ratings completed by an independent assessor who was unaware of treatment allocation, were at pretreatment/wait list, post-*

Clark 1999 *(Continued)*

		treatment/wait list, 3-month post-treatment follow-up, and 12-month post-treatment follow-up."
Incomplete outcome data (attrition bias) Short-term	Low risk	*"43 patients were randomized. One patient (allocated to FCT) dropped out after one session, having indicated that she was much improved and could not arrange time off work for further sessions. All other patients completed treatment"* (and assessments). The proportion of missing outcomes is not enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"During the 1980s, several effective cognitive—behavioral treatments for panic disorder were developed. The two that have been most extensively evaluated are the panic control treatment (PCT) developed by Barlow, Craske, and colleagues and the cognitive therapy program developed by Clark, Salkovskis, Beck, and colleagues."* *"To maximize the amount of change achieved in each therapy session, we developed a set of self-study modules covering the main aspects of therapy and asked patients to complete the modules prior to therapy sessions."*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"Regular individual supervision was provided throughout the trial."* Insufficient information provided

Cottraux 2009

Methods	**Study design:** Randomised controlled trial
Participants	**Inclusion criteria:** DSM-IV diagnosis of panic disorder with agoraphobia. Eligible patients were not allowed to take any psychotropic medication, with the exception of low doses hypnotics, and could not receive psychotherapy during the study. **Exclusion criteria:** current major depression, or a score greater than 18 on the Hamilton rating scale for depression; bipolar disorder, schizophrenia or other psychotic disorders; alcoholism, or street drugs use; history of CBT for PDA, or a current psychotherapy; treatment with antidepressants, neuroleptics, anxiolytics or mood stabilisers within the 2 weeks preceding the entry. **Characteristic of the sample:** • **Age:** VRET group mean age 37.7 years (SD 7.3); CBT mean age 36.6 years (SD 10.6); WL mean age 37 (SD 11.3) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** 0%
Interventions	Participants (n = 92) were randomly assigned to either: **1) Virtual reality exposure therapy** (classified as BT, n = 29; n = 43 after re-randomisation of WL patients) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12

Cottraux 2009 *(Continued)*

- **Duration of intervention:** 12 weeks

2) Cognitive behaviour treatment (classified as CBT, n = 31; n = 44 after re-randomisation of WL patients)

- **Therapy format:** not specified
- **Duration of each session:** 60 minutes
- **Mean number of sessions:** 12
- **Duration of intervention:** 12 weeks

3) Wait list (classified as WL, n = 32)

- **Mean number of sessions:** 0
- **Duration of intervention:** 12 weeks

Outcomes	**Time points for assessment:** baseline, post-treatment/WL (3 months), 6 months follow-up, 12 months follow-up **Measures:** Fear Questionnaire (FQ), Panic Disorder Severity Scale (PDSS), Chambless Agoraphobic Cognitions scale (ACQ), Panic, Phobia and Generalized Anxiety Scale (PPGAS), State and Trait Anxiety questionnaire (STAI), Hamilton Anxiety Rating Scale (HARS), 21-item Beck Depression Inventory (BDI), Sheehan Disability Scale (SDS), Global Assessment of Functioning scale (GAF), Dissociative Experience Scale (DES), Work and Social Adjustment scale (WSA) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** at least 50% reduction of FQ-Ag score (original sample) at post-treatment **ST-Dropouts:** non-completers (original sample) **Continuous scale:** Panic Disorder Severity Scale (PDSS; measured on original + re-randomised sample) **LT-Remission/Response:** although measured (at least 50% reduction of FQ-Ag score at 12 months follow-up) data were not entered in the analyses because dropouts exceeded 30% of originally randomised sample (see Secondary outcomes).
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Low risk	*"Randomization was kept secret and delivered by the biostatistics department of the CHU of Lyon through a phone call to the secretary of each center"*
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	At post-treatment, 63 patients (on 92 originally randomised) were assessed (see study flow chart).
Incomplete outcome data (attrition bias) Long-term	High risk	At 12 months follow-up, 51 patients (on 87 randomised) were assessed (see study flow chart).

Cottraux 2009 *(Continued)*

Selective reporting (reporting bias)	Low risk	Study protocol available. All of the study's pre-specified (primary and secondary) outcomes that are of interest in the review have been reported in the pre-specified way.
Researcher allegiance	Unclear risk	Insufficient information provided
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Craske 1995

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 65 years, principal diagnosis of panic disorder with or without agoraphobia according to DSM-III-R criteria, willingness to random assignment to 17 weeks of either placebo or varying dosage regimes of a psychoactive medication, successful withdrawal from psychotropic medications for at least 7 days prior to initial diagnostic evaluation. **Exclusion criteria:** hypersensitivity to benzodiazepines; diagnoses of organic disorders, obsessive-compulsive disorder, psychoses, bipolar disorder, adjustment disorder and current (within the last 6 months) substance abuse/dependence; suicidality; serious medical conditions. **Characteristic of the sample:** • **Age:** 36.1 years (SD 11, range 21 to 57) • **Percentage of agoraphobic patients:** 67% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 30) were randomly assigned to either: **1) Cognitive behaviour therapy** (classified as CBT, n = 16) • **Therapy format:** individual • **Duration of each session:** 75 minutes • **Mean number of sessions:** 4 • **Duration of intervention:** 4 weeks **2) Nondirective supportive therapy** (classified as SP, n = 14) • **Therapy format:** individual • **Duration of each session:** 75 minutes • **Mean number of sessions:** 4 • **Duration of intervention:** 4 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** panic disorder and agoraphobia sections of the ADIS-R, Anxiety Sensitivity Index (ASI), Fear Questionnaire (FQ), Four Dimensional Anxiety, Subjective Symptoms Scale The following outcomes were used for quantitative analyses: **ST-Remission:** ASI reduced from baseline and < 28 at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers

Craske 1995 *(Continued)*

	Continuous scale: Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	ASI is a self rated measure
Incomplete outcome data (attrition bias) Short-term	Unclear risk	*"Sixteen subjects were assigned randomly to CBT, and 14 to NST. One subject dropped out from NST, none dropped out from CBT."* The proportion of missing outcomes is not enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Craske MG is author of a CBT manual (see Barlow 2000b)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"Treatment integrity was addressed via manualized treatments, and ongoing therapy supervision, with review and feedback of approximately 25% of audiotapes of treatment sessions by the principal author."* It is unclear whether all sessions were recorded and selection of audiotapes was randomised.

Craske 2005a

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-IV principal diagnosis of panic disorder with or without agoraphobia. Those who were medicated at the time of the initial diagnostic evaluation were withdrawn from psychotropic medications over a minimum of 4 weeks and washed out for at least 2 weeks prior to a repeat diagnostic evaluation to re-determine study eligibility. **Exclusion criteria:** history of bipolar disorder, psychosis, posttraumatic stress disorder or current substance abuse/dependence **Characteristic of the sample:** • **Age:** not specified for randomised sample • **Percentage of agoraphobic patients:** 29.2% • **Percentage of patients on drug therapy:** 0%

Craske 2005a *(Continued)*

- **Percentage of patients with major depression:** 30.3%

Interventions	Participants (n = 43) were randomly assigned to either: **1) Cognitive behaviour therapy** (classified as CBT, n = 27) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 11 • **Duration of intervention:** 10 weeks **2) Wait list** (classified as WL, n = 16) • **Mean number of sessions:** 0 • **Duration of intervention:** 10 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 12 months (after commencement) follow-up **Measures:** panic disorder severity (ADIS-IV), Anxiety Sensitivity Index (ASI), Fear Questionnaire-Agoraphobia subscale (FQ-Ag), Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), Subjective Symptoms Scale (SSS) The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	* *"Defined as zero daytime and zero nocturnal panic attacks per week over last 2 weeks, panic disorder severity of 3 or less, and no/mild agoraphobia"*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Unclear risk	*"Of 43 participants, 27 were assigned to CBT; 3 (11.1%) withdrew during treatment. Sixteen were assigned to WL; none withdrew during WL, but 3 (18.8%) withdrew before (n = 2) or during (n = 1) delayed CBT. Reasons for withdrawal are not known."* It is unclear whether non-completers were assessed at post-treatment (probably so)
Incomplete outcome data (attrition bias)	Unclear risk	Number of assessed patients at 12 months follow-up is not reported

Craske 2005a *(Continued)*
Long-term

Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Craske MG is author of a CBT manual (see Barlow 2000b)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"Each treatment session was audiotaped and 25% (n = 112) were selected randomly for independent adherence ratings of each content item of each session (1 = none, 7 = complete adherence) 2 and percent of off-task discussion. [...] Average adherence ratings ranged from 4.95 (SD = 0.77) to 6.01 (SD = 1.00), with a total average of 5.64 (SD -- 0.96), indicating good adherence overall."*

Creager Berger 2001

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age 18 to 60 years, DSM-IV diagnosis of panic disorder **Exclusion criteria:** exhibiting characteristics of, or currently diagnosed with schizophrenia or any personality disorder; evidence of organic brain syndrome or mental retardation; any change in psychotropic or other medications or currently taking a medication for less than 2 weeks; medical conditions that would interfere with the diagnosis and/or treatment of panic disorder not due to a medical condition; report or exhibition of characteristics of present substance abuse that would meet DSM-IV criteria; unwillingness or inability to give informed consent; experiencing less than 3 panic attacks within a 4-week period **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 35.3 years, SD 10.14) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified on ITT sample (90% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = unclear) were randomly assigned to either: **1) Breathing retraining** (classified as PT, n = unclear) • **Therapy format:** group/individual • **Duration of each session:** 45 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks **2) Cognitive therapy** (classified as CBT, n = unclear) • **Therapy format:** group/individual • **Duration of each session:** 67.5 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks
Outcomes	**Time points for assessment:** pre-treatment, baseline (first session), termination (last session), 4 weeks after termination **Measures:** panic diary, Beck Depression Inventory (BDI), Panic Disorder Severity Scale (PDSS), end tidal carbon dioxide level ($ETCO_2$), respiratory rate

Creager Berger 2001 *(Continued)*

	The following outcomes were used for quantitative analyses: **ST-Remission:** although reported (panic-free at termination), ST-Remission could not be calculated following an ITT principle (number of patients randomised to each arm not specified) **ST- Response:** not measured (not imputed) **ST-Dropouts:** not reported **Continuous scale:** Panic Disorder Severity Scale (PDSS) at pre-treatment and at 4 weeks after termination **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"Random assignment occurred by using the random number table from a Sharp scientific calculator"*
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"Of those accepted into the study, one became ineligible after beginning antidepressants after session 3 of the cognitive therapy group, one failed to attend her 6th breathing retraining session and could not be reached by phone, one dropped out of the cognitive behavioral therapy after session 6, five people failed to show up for the first session and could not be reached by phone, three people stated that they were too busy or not interested in beginning the study, and four people did not return phone calls after the initial screening."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"The primary investigator of the study trained each of the therapists and met with each therapist for 2 hours per week during the treatment phase to review the previous session and preview the following session. Additionally, the CBT therapists brought a script into each session with them in order to ensure thorough deliverance of the treatment."*

De Ruiter 1989

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of Panic Disorder with Agoraphobia. Only patients who recognised the symptoms induced by voluntary hyperventilation as similar to their panic attacks were included in the study. **Exclusion criteria:** psychotic symptoms; substance abuse

De Ruiter 1989 *(Continued)*

	Characteristic of the sample: • **Age:** mean 34 years (SD 9.2, range 22 to 60). These data probably refer to the completers sub-sample. • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified for ITT sample (49% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 49) were randomly assigned to either: **1) Breathing retraining/cognitive restructuring** (classified as CBT, n = 17) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks **2) Exposure therapy** (classified as BT, n = 17) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks **3) Breathing retraining/cognitive restructuring + exposure therapy** (classified as CBT, n = 15) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks
Outcomes	**Time points for assessment:** baseline (4 weeks prior to treatment), pre-treatment, post-treatment **Measures:** Fear Surrey Schedule-III (FSS-IZZ), phobic anxiety and avoidance scales, panic attack diary, Bodily Sensations Questionnaire (BSQ), Symptom Checklist-90 (SCL-90), respiratory rate (RR) and end tidal carbon dioxide pressure ($pC0_2$). The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Bodily Sensations Questionnaire (BSQ) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided

De Ruiter 1989 *(Continued)*

Blinding of outcome assessment (detection bias) ST-Remission	High risk	BSQ (used to impute ST-Remission) is a self rated measure
Incomplete outcome data (attrition bias) Short-term	High risk	*"Of the 49 patients who entered treatment, 40 completed the program. Attrition rates were 4 (24%) for BRCR, 4 (24%) for EXP and 1 (6%) for BRCR + EXP"*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"Supervision by a senior clinical psychologist (the second author) was provided on a weekly basis."* Unclear whether all sessions for all patients were supervised.

Dow 2000

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** between 18 and 60 years of age; DSM-IV diagnosis of panic disorder, with or without agoraphobia; current episode duration of at least 3 months; consider panic the main problem; willing to accept random allocation, including the wait list condition. All patients taking medication at the time of entry must have been on a stable dose for 3 months and must have been willing and able to remain on a stable regime for 3 months during the course of treatment. **Exclusion criteria:** depressive disorder severe enough to require urgent treatment; undergoing CBT for the current episode; evidence of organic mental disorder, schizophrenia, alcohol or drug dependence, cardiovascular disease, asthma, epilepsy, or pregnancy or intention to become pregnant during the course of the study. Concurrent Axis II personality disorder was not a reason for exclusion unless the personality disorder was clearly the primary problem. **Characteristic of the sample:** • **Age:** mean age 36.8 years (SD 10) • **Percentage of agoraphobic patients:** 76.1% • **Percentage of patients on drug therapy:** not specified for ITT sample (among completers: 46.4% of Australian patients and 14.1% of Scottish patients) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 186) were randomly assigned to either: **1) CBT - 12 sessions - therapist delivered** (classified as CBT, n = 45) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) CBT - 6 sessions - therapist delivered** (classified as CBT, n = 45) • **Therapy format:** individual • **Duration of each session:** 60 minutes

Dow 2000 *(Continued)*

- **Mean number of sessions:** 6
- **Duration of intervention:** 6 weeks

3) CBT - 6 sessions - computer augmented (classified as CBT, n = 50)

- **Therapy format:** individual
- **Duration of each session:** 60 minutes
- **Mean number of sessions:** 6
- **Duration of intervention:** 6 weeks

4) Wait list (classified as WL, n = 46)

- **Mean number of sessions:** 0
- **Duration of intervention:** unclear

Outcomes	**Time points for assessment:** baseline, post-treatment, 6 months follow-up **Measures:** panic frequency, panic-related distress/disability, Fear Questionnaire (FQ), Mobility Inventory for Agoraphobia (MI), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), State–Trait Anxiety Inventory, Trait subscale (STAI–T), Beck Depression Inventory (BDI), Medical Outcomes Survey Short Form 36 (SF-36) The following outcomes were used for quantitative analyses: **ST-Remission:** panic-free at post-treatment **ST- Response:** not measured (not imputed) **ST-Dropouts:** not measured (the number of non-completers for each arm is not specified) **Continuous scale:** although measured, data cannot be used because number of assessed patients is not reported **LT-Remission/Response:** although measured (panic-free at follow-up), data cannot be used because re-randomisation of WL patients leaves only 3 arms, all classified as CBT (comparison not feasible)
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"In total, 186 patients met entry criteria and were offered and accepted a place in the study. Of these, 163 patients (87.6%) commenced treatment (wait list, n=41; CBT6, n=39; CBT6-CA, n=41; CBT12, n=42). Twenty-three patients (14.1%) failed to receive at least three sessions of their respective course of treatment or to provide adequate data and were classified as dropouts."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable

Dow 2000 *(Continued)*

Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All sessions for all treatment conditions were tape-recorded, and a random selection (20%) of tapes were exchanged between sites and rated by Justin A. Kenardy and Michael G. T. Dow to ensure adherence to treatment protocols and therapeutic competence. There were no significant effects for site, treatment, or Site Treatment on protocol adherence or therapeutic adequacy. Therapists also completed a separate checklist for each therapy session to evaluate adherence to the protocol. The correlation between therapist-rated and externally rated protocol adherence was 0.92 (p .001). No significant differences were found on therapist-rated treatment protocol compliance across site or treatment or for Site Treatment. Overall, there was 97.1% protocol adherence."*
Other bias	Unclear risk	Modification of the original sample with replacements. Number and randomisation of replacements not specified.

Dreessen 1994

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 17 and 70 years, primary DSM-III-R diagnosis of panic disorder with no or mild agoraphobic avoidance, panic disorder for at least 12 months, at least a mean of 1 panic attack per week during the previous 4 weeks, asking for treatment of panic disorder, no use of serotonergic antidepressants or benzodiazepines (for at least 4 weeks). **Exclusion criteria:** depressive disorder preceding the current episode of panic disorder or requiring immediate treatment; behaviour therapy received for panic disorder; evidence of organic mental disorders, psychotic disorders, alcohol or drug dependence, cardiovascular disease, asthma, epilepsy; medical contraindication for exposure, behavioural experiments or hyperventilation **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 34.1 years, range 21 to 52) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified (mood disorder 11% among completers)
Interventions	Participants (n = 37) were randomly assigned to either: **1) Cognitive therapy** (classified as CT, n = unclear) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 15 • **Duration of intervention:** 13 weeks (plus 2 sessions after 1 and 6 months) **2) Applied relaxation** (classified as PT, n = unclear) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 15 • **Duration of intervention:** 13 weeks (plus 2 sessions after 1 and 6 months)

Dreessen 1994 *(Continued)*

	Note that *"after the last patient entered the study, a waiting-list control group was formed. The first 11 men and 7 women meeting the same criteria as used for the treatment group were drawn from the waiting-list to form a control group."* This control group is not considered for this review because it is not randomised.
Outcomes	**Time points for assessment:** baseline, post-treatment (13 weeks), 1-month follow-up, 6-month follow-up **Measures:** panic diary, Fear of Fear Questionnaire, Fear Questionnaire (FQ), State–Trait Anxiety Inventory (STAI), Depressive Symptoms Inventory (DSI), Symptom Check List (SCL-90) The following outcomes were used for quantitative analyses: **ST-Remission:** although measured (panic-free at post-treatment) detailed data are not reported **ST- Response:** not measured **ST-Dropouts:** detailed data are not reported **Continuous scale:** although measured, detailed data are not reported **LT-Remission/Response:** although measured (panic-free at 6 months follow-up) detailed data are not reported
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Selective reporting (reporting bias)	High risk	Study protocol unavailable. Pre-planned measures are not reported with sufficient details.
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"Weekly supervision was given by the first author during the whole treatment."* It is unclear whether all sessions for all patients were supervised.

Emmelkamp 1986

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III diagnosis of agoraphobia **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 36 years, SD 18 to 56) • **Percentage of agoraphobic patients:** 100%

Emmelkamp 1986 *(Continued)*

- **Percentage of patients on drug therapy:** not specified for ITT sample (34.8% among completers)
- **Percentage of patients with major depression:** not specified

Interventions	In a first phase of the study, participants (n = 51) were randomly assigned to either: **1) Exposure in vivo** (classified as BT, n = unclear) • **Therapy format:** group • **Duration of each session:** 150 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 3 weeks (1st phase) **2) Rational emotive therapy** (classified as CT, n = unclear) • **Therapy format:** group • **Duration of each session:** 150 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 3 weeks (1st phase) **3) Self instructional training** (classified as CT, n = unclear) • **Therapy format:** group • **Duration of each session:** 150 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 3 weeks (1st phase) In a second phase of the study, all patents received 6 group sessions (150 minutes each) of exposure in vivo.
Outcomes	**Time points for assessment:** baseline, post-test **Measures:** behavioural walk, phobic anxiety and avoidance scales, Fear Questionnaire (FQ), Irrational Belief Test (IBT), Symptom Check List (SCL-90) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** not reported **Continuous scale:** although measured, detailed data are not reported **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Selective reporting (reporting bias)	High risk	Study protocol unavailable. Pre-planned measures are not reported with sufficient details.

Emmelkamp 1986 *(Continued)*

Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"Therapists received a special training for the research project and extensive manuals were used. Therapists were supervised by the senior author."* It is unclear whether all sessions for all patients were supervised.

Erickson 2003

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-IV diagnosis of either panic disorder with or without agoraphobia, OCD, social phobia, generalised anxiety disorder, specific phobia or PTSD. There were no limitations on past or concurrent treatments. **Exclusion criteria:** active substance abuse or dependence or psychosis **Characteristic of the sample (PDA sub-sample):** • **Age:** not specified for PDA sub-sample • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified for PDA sub-sample • **Percentage of patients with major depression:** not specified for PDA sub-sample
Interventions	Participants (PDA sub-sample, n = 36) were randomly assigned to either: **1) Cognitive behaviour therapy** (classified as CBT, n = unclear) • **Therapy format:** group • **Duration of each session:** 120 minutes • **Mean number of sessions:** 11 • **Duration of intervention:** 11 weeks **2) Wait list** (classified as WL, n = unclear) • **Mean number of sessions:** 0 • **Duration of intervention:** 11 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** Beck Anxiety Inventory (BAI) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** although measured (40% reduction in BAI score), data for the PDA sub-sample are not reported **ST-Dropouts:** not reported for PDA sub-sample **Continuous scale:** the only available measure (BAI) is not considered among outcomes of interest for this review (see Secondary outcomes) **LT-Remission/Response:** not measured

Erickson 2003 *(Continued)*

Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"We assessed clinicians' adherence to the protocol by asking independent raters to listen to audiotapes of a random sample of sessions (33 tapes, or one of every four sessions). Overall, the ratings indicated that group leaders adhered closely to the intended protocol and that quality of implementation was midway between good and very good."*

Gloster 2010

Methods	**Study design:** multicentre randomised controlled trial
Participants	**Inclusion criteria:** 18 to 65 years old (see protocol), DSM–IV–TR diagnosis of panic disorder with agoraphobia, score of 18 or more on the Hamilton Anxiety Scale (HAM-A), score of 4 or more on the Clinical Global Impression (CGI). Patients had to agree to discontinue all psychopharmacological medication and were not allowed to have any concomitant psychotherapy. Patients on psychopharmacological medication underwent a washout period prior to baseline. **Exclusion criteria:** unable to comply with the study schedule or requirements; clinically significant suicidal intent; DSM-IV-TR diagnosis of any psychotic or bipolar disorder, borderline personality disorder, or current alcohol dependence; medical condition that could explain symptoms. Other current comorbid diagnoses, including unipolar depression and other anxiety disorders, were allowed unless they were of primary clinical concern. **Characteristic of the sample:** • **Age:** in the 3 groups age mean (SD) were respectively: 35.5 (SD 11); 35.5 (SD 10.4); 35.6 (SD 11.2) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** 43.2%
Interventions	Participants (n = 369) were randomly assigned to either: **1) CBT variant with therapist-guided exposure outside the therapy room** (classified as BT, n = 163) • **Therapy format:** individual • **Duration of each session:** 100 minutes • **Mean number of sessions:** 14

Gloster 2010 *(Continued)*

- **Duration of intervention:** unclear (12 sessions over 6 weeks + 2 booster sessions at unspecified time)

2) CBT variant with non-therapist-guided exposure outside the therapy room (classified as BT, n = 138)

- **Therapy format:** individual
- **Duration of each session:** 100 minutes
- **Mean number of sessions:** 14
- **Duration of intervention:** unclear (12 sessions over 6 weeks + 2 booster sessions at unspecified time)

3) Wait list (classified as WL, n = 68)

- **Mean number of sessions:** 0
- **Duration of intervention:** not specified

Outcomes	**Time points for assessment:** baseline, intermediate (after the 4th session), post-treatment, 6 months follow-up **Measures:** Structured Interview Guide for the Hamilton Anxiety Scale (HAM-A/SIGH-A), Clinical Global Impression (CGI), Panic Agoraphobia Scale (PAS), Mobility Inventory - Agoraphobia subscale (MI-Ag) The following outcomes were used for quantitative analyses: **ST-Remission:** PAS score ≤ 8 at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Panic Agoraphobia Scale (PAS) **LT-Remission/Response:** although measured (PAS score ≤ 8 at 6 months follow-up), data cannot be used because re-randomisation of WL patients leaves only 2 arms, both classified as BT (comparison not feasible)
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Low risk	*"The randomization list was generated at the clinical coordination center (Dresden) by personnel not associated with patient care. The study centers were blind to the assignment of subsequent cases and were informed of treatment status only after a fax documenting the included patient was sent to the clinical coordination center. More numbers for each center were drawn than necessary so that treatment condition of final patients in each study center remained unpredictable, thereby ensuring blinding of the randomization throughout the study."*
Blinding of outcome assessment (detection bias) ST-Remission	High risk	The Panic Agoraphobia Scale (PAS, used to determine ST-Remission) is a self report questionnaire
Incomplete outcome data (attrition bias) Short-term	High risk	A total of 63 patients (34 + 25 + 4) were lost at post-treatment assessment (see study flow chart)

Gloster 2010 *(Continued)*

Selective reporting (reporting bias)	High risk	Not all of the study's pre-specified primary outcomes have been reported (Aggregated Panic Disorder Scale and Mobility Inventory: PDS-MI score). One reported outcome was not pre-specified (PAS).
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All treatment sessions were videotaped, and a randomly selected sample of almost 18% was evaluated. All raters were blind to treatment condition and evaluated each tape using the therapist adherence and competence rating scale for PD and AG. Adherence and competence were assessed on the basis of a 9-point scale from 0 (nonexistent) to 8 (optimal adherence/excellent competence).The mean overall ratings of therapy adherence and competence across all sessions were 5.53 (SD 1.29) and 5.73 (SD 1.26), respectively, indicating that therapists demonstrated good levels of adherence to the manual and implemented it with good levels of competence."*

Goldstein 2000

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 65 years, DSM-IV diagnosis of panic disorder with agoraphobia at least 1 year's duration, agoraphobic avoidance at least moderately severe for the prior 6 months. Participants excluded on the basis of recent medication changes were eligible for reconsideration once medications were stabilised in appropriate limits. **Exclusion criteria:** being in therapy elsewhere if not willing to suspend that treatment until the end of the study; on dosages of alprazolam in excess of 1.5 mg daily (or similar dosages for other benzodiazepines); taking antidepressant or antianxiety medication for less than 6 months or change of medication within the last 12 weeks; comorbid diagnoses of thought disorder, major depression, bipolar disorder, or substance dependence; presence of another anxiety disorder more severe than the PDA; DSM-IV diagnosis of any of the following Axis II disorders: paranoid, schizoid, schizotypal, antisocial or borderline **Characteristic of the sample:** • **Age:** mean 38.16 years (range 22 to 63) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** 45.65% • **Percentage of patients with major depression:** 0%
Interventions	Participants (n = 46) were randomly assigned to either: **1) EMDR** (not included in this review, n = 18) • **Therapy format:** not specified • **Duration of each session:** 90 minutes • **Mean number of sessions:** 6 • **Duration of intervention:** 4 weeks **2) Association and relaxation therapy** (classified as PT, n = 13) • **Therapy format:** not specified • **Duration of each session:** 90 minutes • **Mean number of sessions:** 6

Goldstein 2000 *(Continued)*

- **Duration of intervention:** 4 weeks

3) Wait list (classified as WL, n = 15)

- **Mean number of sessions:** 0
- **Duration of intervention:** 4 weeks

Outcomes	**Time points for assessment:** baseline, termination, 5 to 6 weeks after termination **Measures:** Agoraphobic Cognitions Questionnaire (ACQ), Body Sensations Questionnaire (BSQ), Brief Body Sensations Interpretation Questionnaire (BBSIQ), Panic Appraisal Inventory (PAI), Mobility Inventory (MI), Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), Brief Symptom Inventory (BSI), Social Adjustment Scale-Self-Report (SAS-SR), Distress Questionnaire, Panic Disorder Symptom Severity interview (PDSS), panic/anxiety diary The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers **Continuous scale:** although measured, data concerning ART group are not reported. EMDR is not an included treatment, therefore only data on WL patients could be extracted (comparison not feasible) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"Of the 46 participants who entered the study, 4 dropped out prior to the completion of treatment. One dropped out during the waiting list period before she provided posttest data or received her treatment condition assignment. Three participants (one of whom had previously been in the waiting list condition) dropped out or were terminated during EMDR: one because of a marital crisis, another because of deterioration, and a third for repeated cancellations of appointments."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable. Data concerning ART group are not reported.
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"To ensure that therapists adhered to the treatment protocol, all sessions were audio- Or videotaped and reviewed by Alan J. Goldstein prior to supervision meetings, which were held weekly to discuss clinical issues and proper provision of treatment. Two of the authors, Dianne L. Chambless and Kimberly A. Wilson, and their trained research assistants followed detailed integrity checklists that*

Goldstein 2000 *(Continued)*

		assessed adherence to treatment protocol, presence of therapist support and reinforcement, and protocol violations, which included introducing other treatments into the session. Adherence checks were conducted on 31% (n = 80) of all sessions. Of these, 33 were independently rated by additional coders to assess reliability. Average percent agreement was 95% for the integrity items identified a priori to be most important. The adherence monitoring team was not otherwise involved in participants' treatment and was unaware of participants' treatment outcome."
Other bias	Unclear risk	Modification of the original sample with replacements. Number and randomisation of replacements not specified.

Gould 1993

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with or without agoraphobia. Subjects were dissuaded from participating in other therapy or self help procedures during the study. Subjects taking medication for anxiety or depression were allowed to participate if they had been stabilised on the medication for at least 4 weeks and continued to have panic symptoms. **Exclusion criteria:** seizure disorder, kidney disease, stroke, schizophrenia, organic brain syndrome, emphysema, heart attack, chronic hypertension **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 35.7, SD 10.2, range 19 to 59 years) • **Percentage of agoraphobic patients:** not specified for ITT sample (94% among completers) • **Percentage of patients on drug therapy:** not specified for ITT sample (16% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 33) were randomly assigned to either: **1) Bibliotherapy** (not included in this review, n = 12) • **Therapy format:** self help • **Duration of each session:** 0 minutes • **Mean number of sessions:** 0 • **Duration of intervention:** 4 weeks **2) Guided imaginal coping** (classified as CBT, n = 9) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 4 weeks **3) Wait list** (classified as WL, n = 12) • **Mean number of sessions:** 0 • **Duration of intervention:** 7 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** Daily Panic Attack Records (DPAR), Panic Cognitions Questionnaire (PACQ), Panic Symptoms Questionnaire (PASQ), Mobility Inventory for Agoraphobia (MI), Anxiety Sensitivity Index (ASI), Beck Depression Inventory (BDI), Likelihood of Having a Panic Attack, Your Thoughts During a Panic Attack, Coping with Panic Attacks, Panic Self-Efficacy Questionnaire.

Gould 1993 *(Continued)*

	The following outcomes were used for quantitative analyses: **ST-Remission:** panic-free **ST- Response:** panic-free OR 50% reduction in number of panic attacks, panic symptoms **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI). Note that reported SDs are uncommonly low, so we considered them as being SEs. **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Panic frequency (used to determine ST-Remission) was self rated
Incomplete outcome data (attrition bias) Short-term	Low risk	*"Only 2 subjects dropped out: 1 from the WL group and 1 from the BT group. The subject from the WL group reported that she had to move to a different state to seek employment and could no longer continue in the study. The subject from the BT condition completed all the dependent measures of the study, but was not included in the analyses because she had failed to read the book."* For this review we are not considering the bibliotherapy arm, therefore data are missing for only one subject (WL) and reasons for missing outcome data are unlikely to be related to true outcome.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	In bibliotherapy arm, *"subjects read the book Coping with Panic (Clum, 1990)."* In Guided Imaginal Coping arm, *"the protocol for each of the eight treatment sessions was outlined for therapists. These plans were derived primarily from material in Coping with Panic."* Clum G is among the study authors.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"The second author supervised the four therapists weekly in order to ensure the uniformity of treatment procedures, and the research team met regularly for discussion. In addition, some treatment sessions were observed directly, or videotaped and later observed by the first and second authors. A random sample of 8 sessions was observed to ensure treatment integrity."*

Griegel 1995

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with or without agoraphobia. Subjects using psychotropic medications were required to maintain a stable dosage for at least 1 month and throughout treatment and evaluation. **Exclusion criteria:** comorbid DSM-III-R diagnosis rated as severe as panic disorder; being in psychotherapy for anxiety (subjects who were in psychotherapy for other psychological difficulties were included if they had been in a stable therapeutic relationship for at least 3 months). **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 36.03 years, SD 8.02, range 21 to 52) • **Percentage of agoraphobic patients:** unclear (data are measured on 36 patients, not 37. Among these 36 patients, 86% are agoraphobic) • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 37) were randomly assigned to either: **1) Breathing retraining - slow respiration rate** (classified as PT, n = 11) • **Therapy format:** not specified • **Duration of each session:** 42.5 minutes • **Mean number of sessions:** 3 • **Duration of intervention:** 2 weeks **2) Breathing retraining - increase respiration rate** (classified as PT, n = 12) • **Therapy format:** not specified • **Duration of each session:** 42.5 minutes • **Mean number of sessions:** 3 • **Duration of intervention:** 2 weeks **3) Wait list** (classified as WL, n = 14) • **Mean number of sessions:** 0 • **Duration of intervention:** 4
Outcomes	**Time points for assessment:** baseline, post-treatment. **Measures:** Anxiety Disorders Interview Schedule–R (ADIS–R), Anxiety Sensitivity Index (ASI), Emotional Control Questionnaire (ECQ), Self-Efficacy Questionnaire (SEQ), Diagnostic Symptom Questionnaire (DSQ), Interoceptive Exposure Test (IET), respiratory rate The following outcomes were used for quantitative analyses: **ST-Remission:** panic-free **ST- Response:** not measured **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	None

Griegel 1995 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	Panic frequency was rated by assessors. *"Assessors who conducted evaluations at pre- and post-assessment were blind to the subjects' experimental condition."*
Incomplete outcome data (attrition bias) Short-term	High risk	*"The first 37 subjects agreeing to participate were randomly assigned to one of three groups to enlist the requisite sample size of 10 subjects per cell (7 subjects dropped out of the study before completion of the post-assessment)."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"All sessions were audiotaped and 25% of these tapes were rated by an assessor blind to the treatment conditions, who was then asked to identify the treatment condition being conducted. Accurate identification of treatment protocol was made 93% of the time (14/15). One tape was inaudible and could not be rated."* It is unclear whether the selection of tapes to assess was random.

Hazen 1996

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** 18 years of age or older, primary DSM-III-R diagnosis of panic disorder with or without agoraphobia, minimum Grade 8 reading and writing ability, physician agreement regarding participation **Exclusion criteria:** presence of organic disease which might be related to panic disorder or interfere with participation in the study; presence of other serious psychiatric disorders, specifically psychotic disorders, substance abuse and current major depressive disorder; presence of significant suicidal risk; involvement in other psychological treatment; current pharmacological treatment for panic disorder, with the exception of low doses of benzodiazepines (equivalent of 20 mg diazepam or less) or stable doses of antidepressants (i.e. prescribed for at least 6 months and stable dose for at least 3 months) **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 37.12 years, SD 9.57, range 20 to 73 years) • **Percentage of agoraphobic patients:** not specified for ITT sample (92.4% among completers) • **Percentage of patients on drug therapy:** not specified for ITT sample (45% among completers) • **Percentage of patients with major depression:** 0%
Interventions	Participants (n = 117) were randomly assigned to either:

Hazen 1996 *(Continued)*

1) Individual self administration of the self help manual (not included, n = unclear)

- **Therapy format:** individual
- **Mean number of sessions:** 0
- **Duration of intervention:** weeks

2) Use of the manual in a self help treatment group (not included, n = unclear)

- **Therapy format:** group
- **Duration of each session:** 90 minutes
- **Mean number of sessions:** 13
- **Duration of intervention:** 14 weeks

3) Use of the manual in a treatment group led by professional therapists (classified as CBT, n = unclear)

- **Therapy format:** group
- **Duration of each session:** 90 minutes
- **Mean number of sessions:** 13
- **Duration of intervention:** 14 weeks

4) Wait list (classified as WL, n = unclear)

- **Mean number of sessions:** 0
- **Duration of intervention:** 14 weeks

Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** Anxiety Sensitivity Index (ASI), Fear Questionnaire-Agoraphobia Subscale (FQ-Ag), Sheehan Patient-Rated Anxiety Scale (SPRAS), Clinical Global Improvement (CGI) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (not imputed: number of randomised patients for each arm not specified) **ST- Response:** not measured (not imputed: number of randomised patients for each arm not specified) **ST-Dropouts:** not reported by treatment group **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"Of the 117 subjects enrolled in the evaluation study, 106 completed the Anxiety Sensitivity Index at pre- and posttreatment. These subjects comprised the sample for the present study."* No further detail about these 11 dropouts is reported.

Hazen 1996 *(Continued)*

Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Hendriks 2010

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** adults aged over 60 years, principal DSM-IV diagnosis of panic disorder with or without agoraphobia. Co-morbidity with other anxiety disorders, depression or dysthymia was allowed as long as PD(A) was the principal diagnosis. Participants using benzodiazepines were asked to adhere to a fixed daily dose for the duration of the study. **Exclusion criteria:** presence of severe psychiatric disorders (e.g. psychotic disorder, bipolar disorder), severe somatic condition that would hinder appropriate application of CBT (e.g. severe cardiovascular disease), contraindication for paroxetine, current use of an antidepressant in an adequate dose, current and adequate psychological treatment, failure of paroxetine or CBT in the past, abuse of or dependency on alcohol or psychoactive substances, dementia and a score of 23 or less on the Mini-Mental State Examination **Characteristic of the sample:** • **Age:** mean age 68.6 years (SD 4.6) • **Percentage of agoraphobic patients:** 48% • **Percentage of patients on drug therapy:** 22% • **Percentage of patients with major depression:** not specified (mood disorder 12.2%)
Interventions	Participants (n = 49) were randomly assigned to either: **1) Cognitive behaviour therapy** (classified as CBT, n = 20) • **Therapy format:** individual • **Duration of each session:** 50 minutes • **Mean number of sessions:** 14 • **Duration of intervention:** 14 weeks **2) Paroxetine** (not included, n = 17) • **Therapy format:** individual • **Duration of each session:** 30 minutes • **Mean number of sessions:** 9 • **Duration of intervention:** 14 weeks **3) Wait list** (classified as WL, n = 12) • **Mean number of sessions:** 0 • **Duration of intervention:** 14 weeks
Outcomes	**Time points for assessment:** baseline, 8 weeks, 14 weeks (termination), 26 weeks (3 months follow-up)

Hendriks 2010 *(Continued)*

Measures: Agoraphobic Cognitions Questionnaire (ACQ), Mobility Inventory (MI), Symptom Checklist (SCL-90)

The following outcomes were used for quantitative analyses:

ST-Remission: panic-free (zero panic attacks in the preceding week) at termination

ST- Response: improvement > 30% on one of the primary outcome scales

ST-Dropouts: non-completers

Continuous scale: Agoraphobic Cognitions Questionnaire (ACQ)

LT-Remission/Response: although measured (panic-free at 26 weeks), details are not reported and anyway refer to the comparison paroxetine versus CBT

Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"A sealed envelope was randomly selected from an initial total of 75 envelopes containing the treatment assignments, with 30 being labelled as CBT, 30 as paroxetine and 15 as waiting list."*
Allocation concealment (selection bias)	Unclear risk	See above. It is unclear whether envelopes were opaque and sequentially numbered
Blinding of outcome assessment (detection bias) ST-Remission	High risk	*"All assessments were administered by trained, independent psychologists who were blind to the study and treatments delivered."* However, panic frequency was rated through MI, which is a self rated measure.
Incomplete outcome data (attrition bias) Short-term	Low risk	*"Five patients (10.2%) failed to complete the 14-week treatment protocol: three dropped out in the paroxetine condition (side-effects, n = 1; protocol violation, n = 1; broken hip, n = 1), one in the CBT condition (protocol violation) and one in the WL condition (severe somatic illness."* We are not considering the paroxetine group for this review, therefore the proportion of missing outcomes compared with the observed event risk is not enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Low risk	Study protocol available. All primary outcomes coincide (however, note that BDI, a secondary outcome, is not reported).
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"Throughout the study, they were weekly supervised by a registered supervisor (a member of the Dutch Association of Behavioural and Cognitive Therapy). Per session, the therapists recorded which specific CBT component they had applied and any deviations from the treatment manual were discussed."* It is unclear whether all sessions for all patients were supervised; no detail is reported about supervisor assessments.
Other bias	Unclear risk	This study has been funded by Glaxo Smith Kline, so a sponsorship bias is possible. However, because the Paroxetine arm is not considered for this review, it is unclear whether the comparison CBT versus WL can be affected by this possible source of bias.

Hoffart 1995

Methods	**Study design:** randomised controlled trial, inpatient setting
Participants	**Inclusion criteria:** age from 20 to 65 years, DSM-III-R diagnosis of panic disorder with agoraphobia, DSM-III-R agoraphobia severity rated as moderate or severe, patients considered the symptoms of agoraphobia (that is avoidance behaviour and situational panic or symptom attacks and not spontaneous panics or other mental problems) as their main problem. A plan for the reduction or discontinuation of medication before hospital admission was agreed upon and patients were informed that use of psychotropic medication was prohibited during the 6-week inpatient treatment period. **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 40.1 years, SD 9.3) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified for ITT sample (63% among completers)
Interventions	Participants (n = 52) were randomly assigned to either: **1) Cognitive therapy** (classified as CBT, n = 26) • **Therapy format:** group • **Duration of each session:** unclear • **Mean number of sessions:** unclear • **Duration of intervention:** 6 weeks **2) Guided mastery therapy** (classified as BT, n = 26) • **Therapy format:** group • **Duration of each session:** unclear • **Mean number of sessions:** unclear • **Duration of intervention:** 6 weeks
Outcomes	**Time points for assessment:** precare (before hospital admission), pre-treatment (1 week), post-intensive period (4 weeks), post-treatment (6 weeks), 1 year follow-up **Measures:** Behavioral Avoidance Tests (BATs), Structured Clinical Interview for the DSM-III-R (SCID-I), Phobic Avoidance Rating Scale (PARS), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), Mobility Inventory for Agoraphobia (M), Self-Efficacy Scales for Agoraphobia (SESA), State-Trait Anxiety Inventory (STAI), Beck Depression Inventory (BDI), panic diary The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* at post-treatment **ST- Response:** improvement > 50% on PARS separation avoidance sub-scale **ST-Dropouts:** non-completers **Continuous scale:** Body Sensations Questionnaire (BSQ) **LT-Remission/Response:** high end-state functioning* at 1 year follow-up
Notes	* "*It was decided a priori to give a status of high endstate functioning to those who at posttreatment (l) had a score of 1.5 or lower at the PARS separation avoidance subscale, implying that at least half of the six situations of this subscale were approached regularly without use of safety signals (e.g. medication); (2)*

Hoffart 1995 *(Continued)*

had a score of 3--"symptoms interfere with work or social activity only in minor ways"--or less in interviewer rated global severity; and (3) were free of spontaneous panic attacks in the two weeks after discharge."

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	*"A symptom rating interview was performed by a psychiatrist who worked in another institution and was blind to the treatment condition of the patients. The interview included the 0 to 4 point Phobic Avoidance Rating Scale (PARS)."*
Incomplete outcome data (attrition bias) Short-term	Unclear risk	*"Six of the 52 patients, 3 in each condition, dropped out from the study. Four patients dropped out before or within the first 24 hours after admission to the hospital: 2 because of problematic family circumstances, 1 was not allowed financial coverage for the hospital stay from her home country, and 1 withdrew as she experienced overwhelming fantasies of being locked in forever in a mental hospital. Two patients, 1 in each condition, withdrew just after the discontinuation of anxiolytics because they could not tolerate being without them."* Missing outcome data are balanced in numbers across intervention groups, with similar reasons for missing data across groups. However, it is unclear whether the proportion of missing outcomes is enough to have a clinically relevant impact on the intervention effect estimate.
Incomplete outcome data (attrition bias) Long-term	Unclear risk	All treatment completers (n = 46, see above) were assessed at 1-year follow-up
Selective reporting (reporting bias)	High risk	Study protocol unavailable. Results of ACQ measurements are not reported.
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	*"The two psychologists alternated between being supervisor and being the primary therapist, responsible for the overall treatment and for conducting the group sessions together with a co-therapist, in the 6 groups. In the intensive three-week period, the supervising psychologist conducted two 45 min supervision sessions per week, addressing immediate treatment problems and questions about adherence to the manuals. In addition, the supervisor listened to audiotapes of therapy sessions and gave written feedback to the therapists."* It is unclear whether all therapy sessions were assessed.

Karekla 2004

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** primary DSM-IV diagnosis of panic disorder with or without agoraphobia **Exclusion criteria:** psychosis, substance abuse, suicidal ideation or intent

Karekla 2004 *(Continued)*

	Characteristic of the sample: • **Age:** not specified for ITT sample (among those who attended at least the first treatment session, mean age was 34.95, SD 11.07, range 20 to 67) • **Percentage of agoraphobic patients:** not specified for ITT sample (78.3% among those who attended at least the first treatment session) • **Percentage of patients on drug therapy:** not specified for ITT sample (72.7% among those who attended at least the first treatment session) • **Percentage of patients with major depression:** not specified for ITT sample (34.8% among those who attended at least the first treatment session)
Interventions	Participants (n = 28) were randomly assigned to either: **1) Panic control treatment** (classified as CBT, n = 14) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks **2) Acceptance-enhanced panic control treatment** (classified as 3W, n = 14) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks
Outcomes	**Time points for assessment:** pre-treatment, mid-treatment, post-treatment, 6 months follow-up **Measures:** State-Trait Anxiety Inventory (STAI), Acceptance and Action Questionnaire (AAQ), Automatic Thoughts Questionnaire (ATQ), Beck Depression Inventory (BDI), Anxiety Sensitivity Index (ASI), Agoraphobic Cognitions Questionnaire (ACQ), Panic and Agoraphobia Scale (PAS), Panic Disorder Severity Scale (PDSS), quality of life (SF-36), Valued Living Questionnaire (VLQ), White Bear Suppression Inventory (WBSI), Subjective Units of Distress Scale (SUDS), Diagnostic Symptoms Questionnaire (DSQ), Anxiety and Willingness Scale (AWS) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers (including early dropouts) **Continuous scale:** although measured, data cannot be used because number of assessed patients is not reported **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided

Karekla 2004 *(Continued)*

Incomplete outcome data (attrition bias) Short-term	High risk	*"Twenty-eight individuals met inclusion criteria and were scheduled for treatment. Of those, 22 attended at least the first session. Fourteen participants completed the full 10 weeks of treatment. The cases of participants who prematurely dropped out of treatment are dropped out from further analyses."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Klosko 1988

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 65 years, DSM-III-R primary diagnosis of panic disorder with a clinician's severity rating of at least 4 on a 0 to 8 scale (moderate severity), at least 1 panic attack in the week before starting treatment **Exclusion criteria:** pharmacotherapy or psychotherapy begun in the past 6 months; either in drug or psychotherapeutic treatment more than 6 months (unless subjects agreed to stop such treatment for the duration of the study); on 4 mg or more of alprazolam for any 3-week period and were non-responders; evidence of benzodiazepine hypersensitivity; undergone cognitive behaviour therapy for anxiety at any time; females who were pregnant or lactating or at risk to become pregnant; significant medical problems, as determined by history, medical report and laboratory values; history of psychotic disorder or dementia; history of alcohol or other substance abuse within the last 6 months; current or past bipolar disorder. Subjects with major depression were excluded only if depression predominated over panic disorder at the time of presentation and if depression preceded panic disorder chronologically. Subjects with acute suicidal ideation were excluded. **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 37, SD 11.04) • **Percentage of agoraphobic patients:** not specified for ITT sample (79% among completers) • **Percentage of patients on drug therapy:** not specified for ITT sample (59.6% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 69) were randomly assigned to either: **1) Alprazolam** (not included, n = 17) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks **2) Placebo** (not included, n = 18) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks

Klosko 1988 *(Continued)*

	3) Behaviour therapy (classified as CBT, n = 18) • **Therapy format:** individual • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks **4) Wait list** (classified as WL, n = 16) • **Mean number of sessions:** 0 • **Duration of intervention:** 15 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment. **Measures:** daily self monitoring, Anxiety Disorders Interview Schedule - R (ADIS-R), Hamilton Anxiety Rating Scale (HAM-A), Hamilton Rating Scale for Depression (HAM-D) The following outcomes were used for quantitative analyses: **ST-Remission:** ADIS-IV severity < 4 **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Disorders Interview Schedule - R (ADIS-R) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	*"Posttreatment clinical assessment measures were gathered through administration of a short form of the ADIS-R. The ADISR administrators were blind to group assignment"*
Incomplete outcome data (attrition bias) Short-term	High risk	*"Out of 69 initial subjects, 57 subjects completed the study, and 12 subjects dropped out. A higher rate of dropout was observed in the placebo group compared with the other three groups. One subject out of 17 (5.9%) dropped from the alprazolam group, 7 out of 18 (38.9%) from the placebo group, 3 out of 18 (16.7%) from the PCT group, and 1 out of 16 (6.3%) from the waiting-list group."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Barlow DH is author of a CBT manual (see Barlow 2000b)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All therapy sessions were tape-recorded and checked for treatment integrity."*

Klosko 1988 *(Continued)*

Other bias	High risk	Performance bias: in contrast with patients in the CBT group, patients in the WL were not asked to withdraw medications.

Korrelboom 2013

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age 18 to 65 years, DSM-IV-TR clinical diagnosis of panic disorder with or without agoraphobia, being able to read, speak and understand Dutch well enough to participate in the treatment groups and to fill in all assessments. Use of psychotropic medication was permitted only if the medication dose was stable during the study and had been stable for at least 2 months prior to study inclusion. Patients who changed their medication during treatment were considered to be dropouts, patients who started (a new) medication less than 2 months before inclusion had to wait for participation until they had fulfilled this 2-month criterion. **Exclusion criteria:** severe co-morbid psychopathology, such as psychosis; addiction; being suicidal; mental retardation; concurrent psychological treatments, or cognitive behavioural treatments in the past 6 months **Characteristic of the sample:** • **Age:** mean age 36.1 years (SD 11.9) • **Percentage of agoraphobic patients:** 79% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 143) were randomly assigned to either: **1) Competitive memory training for panic** (classified as CBT, n = 70) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 7 • **Duration of intervention:** 7 weeks **2) Applied relaxation** (classified as PT, n = 73) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 7 • **Duration of intervention:** 7 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 6 months follow-up **Measures:** Panic Appraisal Inventory (PAI), Mobility Inventory (MI) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Mobility Inventory for agoraphobia when alone (MI-A) **LT-Remission/Response:** not measured

Korrelboom 2013 *(Continued)*

Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	MI is a self rated scale
Incomplete outcome data (attrition bias) Short-term	Low risk	*"During treatment, a total of 24 patients (17%) dropped out (13 in AR and 11 in COMET). Analyses were on an intention-to-treat basis. First, in a binary logistic regression analysis, it was checked whether dropout was predicted by age, diagnosis, gender or any of the outcome measures at M-pre. Then missing values were imputed with the SPSS 20 multiple imputation algorithm."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	It seems that authors are involved in conceptualisation of COMET-panic: *"Since this COMET-panic protocol appeared to be effective in two pilot studies in routine clinical settings (Korrelboom et al. 2008; Peeters et al. 2005), it was decided to put the new protocol to the test in a randomized controlled trial versus an evidence-based anti-panic procedure, in this case AR."*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"Both panic management techniques were available in written form (a manual) for therapists and patients. In regular meetings therapists and researchers discussed adherence to the treatment protocols. As an additional check on treatment integrity, patients filled in a checklist containing specific questions about the nature of the treatments they had just received. This checklist contained statements about specific differential identifying elements of both therapies. If both treatments had been delivered properly, AR patients should identify more AR ingredients in the treatment they had received and few/none of the COMET elements, and vice versa for the COMET patients. In the AR condition the mean score for 'AR ingredients' was 36.7 (SD = 3.1) whereas the highest possible score was 40, while the mean score for 'COMET ingredients' in this condition was only 16.6 (SD = 9.5) whereas the lowest possible score was 4. On the other hand, in the COMET condition these figures were 34.0 (SD = 5.7) for 'COMET ingredients' (possible maximum of 40) and 15.2 (SD = 7.6) for 'AR ingredients' (possible minimum of 4)."*

Lidren 1994

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with or without agoraphobia. Individuals taking medication for anxiety could participate if they were still suffering from panic symptoms after 6

Lidren 1994 *(Continued)*

	weeks of stabilisation on medication, if they maintained the same dosages throughout treatment, and if they recorded both the type and amount of medication usage throughout the study. **Exclusion criteria:** seizure disorder, kidney disorder, stroke, myocardial infarction, chronic hypertension, emphysema, organic brain syndrome, chronic use of alcohol, drug dependence, major depressive disorder, psychotic disorders, involvement in any type of therapy focusing on anxiety management **Characteristic of the sample:** • **Age:** 33.7 years (SD 11.8) • **Percentage of agoraphobic patients:** 83.3% • **Percentage of patients on drug therapy:** 39% • **Percentage of patients with major depression:** 0%
Interventions	Participants (n = 36) were randomly assigned to either: **1) Bibliotherapy** (not included, n = 12) • **Therapy format:** self help • **Mean number of sessions:** 0 • **Duration of intervention:** 8 weeks **2) Group therapy** (classified as CBT, n = 12) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks **3) Wait list** (classified as WL, n = 12) • **Mean number of sessions:** 0 • **Duration of intervention:** 8 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 3 months follow-up, 6 months follow-up **Measures:** panic frequency and severity, Panic Attack Symptom Questionnaire (PASQ), Panic Attack Cognition Questionnaire (PACQ), Mobility Inventory (MI), Panic Self-Efficacy Questionnaire (PSEQ), Beck Depression Inventory (BDI) The following outcomes were used for quantitative analyses: **ST-Remission:** Mobility Inventory (MI) score < 32 at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Mobility Inventory for agoraphobia when alone (MI-A) **LT-Remission/Response:** Mobility Inventory (MI) score < 32 at 6 months follow-up
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided

Lidren 1994 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Mobility Inventory (MI) is a self rated measure
Incomplete outcome data (attrition bias) Short-term	Low risk	*"Attrition rates were zero for all three conditions"*
Incomplete outcome data (attrition bias) Long-term	Low risk	(see above)
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"The BT condition used Clum's (1990) Coping with panic book. Subjects in the GT condition also used Clum's (1990) Coping with panic text."* Glum GA is among the study authors.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"We videotaped all sessions, with the faculty member on the research team viewing these tapes for treatment integrity."*

Malbos 2011

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-IV diagnosis of panic disorder with agoraphobia **Exclusion criteria:** epilepsy, dissociative or non-dissociative chronic psychosis, recent discontinuation of psychotropic drugs, substance dependence **Characteristic of the sample:** • **Age:** mean age 44.11 years (SD 13.79, range 24 to 72) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** 23.5%
Interventions	Participants (n = 17) were randomly assigned to either: **1) Virtual reality exposure only** (classified as BT, n = 9) • **Therapy format:** not specified • **Duration of each session:** 90 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** 10 weeks **2) Virtual reality exposure + cognitive therapy** (classified as CBT, n = 8) • **Therapy format:** not specified • **Duration of each session:** 90 minutes • **Mean number of sessions:** 10

Malbos 2011 *(Continued)*

	• **Duration of intervention:** 10 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 3 months follow-up **Measures:** Presence Questionnaire (PQ v3.0), Subjective Units of Discomfort (SUD), Depression Anxiety Stress Scales (DASS 21), Anxiety Sensitivity Index (ASI), Agoraphobia Cognitions Questionnaire (ACQ), Mobility Inventory for Agoraphobia (MI), Simulation Sickness Questionnaire (SSQ), Behavioural Avoidance Test (BAT), heart rate (HR) and heart rate variability (HRV) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"The allocation to each group was done using a randomisation table generated by a computerised sequence generator."*
Allocation concealment (selection bias)	Low risk	*"It was impossible to foresee the assignment of the next patient entering the study"* (personal communication)
Blinding of outcome assessment (detection bias) ST-Remission	High risk	ASI is a self rated measure
Incomplete outcome data (attrition bias) Short-term	Low risk	*"One participant* (VRO)*dropped out at an early stage due to a severe myopia."* The proportion of missing outcomes is not enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Unclear risk	*"In the present study we designed all required VEs for the VRET of agoraphobia within a game level editor."* Although possible, the extent to which this source of bias may affect the results is unclear, because both arms are administered the same VEs.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Marchione 1987

Methods	**Study design:** randomised controlled trial

Marchione 1987 *(Continued)*

Participants	**Inclusion criteria:** DSM-III diagnosis of agoraphobia with panic attacks, other inclusion criteria are mentioned but not reported. **Exclusion criteria:** none reported **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 38.1 years, SD 11.8, range 25 to 65) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 19) were randomly assigned to either: **1) Cognitive therapy + graduated exposure** (classified as CBT, n = unclear) • **Therapy format:** group • **Duration of each session:** 120 minutes (first 2 sessions), 60 minutes (following 14 sessions) • **Mean number of sessions:** 16 • **Duration of intervention:** not specified **2) Progressive deep muscle relaxation + graduated exposure** (classified as BT, n = unclear) • **Therapy format:** group • **Duration of each session:** 120 minutes (first 2 sessions), 60 minutes (following 14 sessions) • **Mean number of sessions:** 16 • **Duration of intervention:** not specified **3) Graduated exposure alone** (classified as BT, n = unclear) • **Therapy format:** group • **Duration of each session:** 120 minutes (first 2 sessions), 90 minutes (following 14 sessions) • **Mean number of sessions:** 16 • **Duration of intervention:** not specified
Outcomes	**Time points for assessment:** baseline, mid-treatment, post-treatment **Measures:** Global Assessment of Severity, Self-Rating of Severity, Phobic Anxiety and Avoidance Scale, Fear Survey Schedule, Fear Questionnaire (FQ), Taylor Manifest Anxiety Scale, Panic Scale, Beck Depression Inventory (BDI), Hopkins Symptom Checklist, Subjective Symptom Checklist, Standardised Behavioral Avoidance Course (S-BAC), heart rate The following outcomes were used for quantitative analyses: **ST-Remission:** although measured (high end-state functioning*) detailed data are not reported **ST- Response:** not measured **ST-Dropouts:** detailed data are not reported **Continuous scale:** although measured, detailed data are not reported **LT-Remission/Response:** not measured
Notes	** "Five criteria were used to classify subjects' level of Endstate (low-medium-high) functioning. Subjects were assigned 1 point for each of the following scores: (a) ≤2 on the Global Assessment of Severity; (b) ≤2 on the Self-Rating of Severity; (c) ≤3 on the Phobic Anxiety and Avoidance Scale; (d) Completing the Standardized-Behavioral Avoidance Course; and, (e) ≤3 SUDS (in vivo anxiety). Subjects with 0-1 points = low Endstate functioning; 2-3 points = medium Endstate functioning, and those with 4-5 points were classified as having high Endstate functioning."*

Marchione 1987 *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Selective reporting (reporting bias)	High risk	Study protocol unavailable. Pre-planned measures are not reported with sufficient detail.
Researcher allegiance	High risk	*"The cognitive therapy was developed by Michelson (1984), adapted from Antaki and Brewin (1982), Beck and Emery (1979), Bums (1980), Mckay, Davis and Fanning (1981), and Sank and Shaffer (1984)."* Michelson L is among the study authors.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"To ensure the treatment procedures were administered consistently, weekly meetings were held to discuss and review all treatment sessions. Treatment integrity probes were randomly conducted on 25% of the sessions and revealed consistently high (100%) levels of fidelity."*

Meulenbeek 2008

Methods	**Study design:** multicentre randomised controlled trial
Participants	**Inclusion criteria:** over 18 years of age, subthreshold or mild panic disorder with or without agoraphobia (DSM-IV), defined as having symptoms of panic disorder falling below the cut-off of 13 on the Panic Disorder Severity Scale–Self Report (PDSS–SR). If a participant used medication for anxiety or depression (e.g. benzodiazepines or antidepressants) it was agreed to keep medication use stable during the study period. **Exclusion criteria:** severe panic disorder (PDSS–SR > 12), current psychological treatment for panic disorder-related complaints, presence of other severe mental or social problems, suicidal intention warranting treatment or likely to interfere with participation in the group course as assessed by an experienced psychologist during intake. **Characteristic of the sample:** • **Age:** not specified for clinical sub-sample (among all participants, mean age was 42 years, SD 12.4, range 20 to 75) • **Percentage of agoraphobic patients:** not specified for clinical sub-sample (62% among all participants) • **Percentage of patients on drug therapy:** not specified for clinical sub-sample (38.7% among all participants) • **Percentage of patients with major depression:** not specified
Interventions	Participants (clinical sub-sample, n = 100) were randomly assigned to either: **1) "Don't Panic" intervention** (classified as CBT, n = 50) • **Therapy format:** group • **Duration of each session:** 120 minutes • **Mean number of sessions:** 8

Meulenbeek 2008 *(Continued)*

- **Duration of intervention:** 8 weeks

2) Wait list (classified as WL, n = 50)

- **Mean number of sessions:** 0
- **Duration of intervention:** 12

Outcomes	**Time points for assessment:** baseline, post-intervention, 6 months follow-up **Measures:** Panic Disorder Severity Scale–Self Report (PDSS–SR), Mini International Neuropsychiatric Interview–Plus (MINI–Plus), Mobility Inventory (MI), sub-scale for anxiety of the Hospital Anxiety and Depression Scale (HADS–Anxiety), Beck Depression Inventory (BDI–II) The following outcomes were used for quantitative analyses: **ST-Remission:** PDSS at post-treatment < 1 SD compared to baseline mean and below 4 (cut-off value) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** not reported for clinical sub-sample **Continuous scale:** Panic Disorder Severity Scale–Self Report (PDSS–SR, data for clinical sub-sample available on personal communication) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"A blocked randomisation scheme was used, stratified by mental health centre, subthreshold panic disorder v. mild panic disorder, and by presence v. absence of co-occurring agoraphobia."*
Allocation concealment (selection bias)	Low risk	*"The randomisation took place after administration of the Mini International Neuropsychiatric Interview–Plus (MINI–Plus)8 and was carried out centrally by an independent third party."*
Blinding of outcome assessment (detection bias) ST-Remission	High risk	ST-Remission is calculated from the Self Report of Panic Disorder Severity Scale
Incomplete outcome data (attrition bias) Short-term	Low risk	All randomised patients (clinical sub-sample) were assessed with PDSS-SR at post-treatment (personal communication)
Selective reporting (reporting bias)	High risk	Study protocol available: declared primary outcome is "*incidence of DSM-IV panic disorder.*" In published report: "*we used the PDSS–SR and the MINI–Plus as the primary outcome measures.*"
Researcher allegiance	High risk	*"We developed an early intervention for panic symptoms, called the 'Don't Panic' course"*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided
Other bias	High risk	Performance bias: WL patients are free to make use of other interventions

Meuret 2008

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 60 years, principal DSM-IV diagnosis of panic disorder with or without agoraphobia, no additional psychological treatment until after the 2-month follow-up. If on psychotropic medications, on stable doses for at least 3 months prior to treatment with an agreement not to change dosage at least until after the 2-month follow-up **Exclusion criteria:** evidence of organic mental disorder, suicidality, schizophrenia, alcohol or drug dependence, cardiovascular disease, pulmonary disease, epilepsy or pregnancy **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 41 years, SD 8.9) • **Percentage of agoraphobic patients:** not specified for ITT sample (83.8% among completers) • **Percentage of patients on drug therapy:** not specified for ITT sample (32.4% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 43) were randomly assigned to either: **1) Breathing training therapy** (classified as PT, n = 24) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 5 • **Duration of intervention:** 4 weeks **2) Wait list** (classified as WL, n = 19) • **Mean number of sessions:** 0 • **Duration of intervention:** 4 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment (week 4), 2 months follow-up (week 12), 12 months follow-up **Measures:** Panic Disorder Severity Scale (PDSS), Clinical Global Impression Scale (CGI), Anxiety Sensitivity Index (ASI), Sheehan Disability Scale (SDS), Mobility Inventory for Agoraphobia (MI-AAL), Beck Depression Inventory (BDI), week-by-week changes in end-tidal pCO_2 (mm Hg) and RR (breaths/minute) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Panic Disorder Severity Scale (PDSS) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided

Meuret 2008 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	*"An assessment battery was administered at baseline (week 0), post-treatment (week 4), 2-month follow-up (week 12), and 12-month follow-up (week 53). It included the PDSS (Shear et al., 1997), a clinician-rated scale of PD severity and the Clinical Global Impression Scale (CGI; Guy, 1976) (both assessed by independent raters)."* It is unclear whether raters were blind to treatment allocation.
Incomplete outcome data (attrition bias) Short-term	High risk	35 patients were analysed at post-treatment (see study flow chart)
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"We devised a capnometry-assisted breathing training therapy (BRT) that uses immediate feedback to teach patients how to raise their pCO_2 over a series of training and practice sessions."*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Meuret 2010

Methods	**Study design:** 2-site, randomised controlled trial
Participants	**Inclusion criteria:** minimum age of 18 years, current principal DSM-IV diagnosis (i.e. the disorder presently associated with the greatest life interference) of panic disorder with agoraphobia, being on a stable dose of psychotropic medication for at least 3 months before study initiation (if applicable) and agreement to continue this dose through the 2-month follow-up appointment, agreement not to initiate additional therapy until after the final follow-up appointment. **Exclusion criteria:** indication of a history of bipolar disorder, psychotic disorder or suicidal intention, or current substance abuse or dependence, organic mental disorder, serious unstable medical disease, respiratory illness or seizures **Characteristic of the sample:** • **Age:** not reported for ITT sample (among initiators, mean age was 33.2 years, SD 9.9, range 20 to 57) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not reported for ITT sample (62.2% among initiators) • **Percentage of patients with major depression:** not reported for ITT sample (mood disorders 3.3% among initiators)
Interventions	In a first phase of the study, participants (n = 47) were randomly assigned to 5 weekly sessions of either capnometry-assisted breathing training (n = 24) or cognitive training (n = 23) In a second phase of the study, participants of both groups underwent 3 weekly sessions of in vivo exposure plus a 4th session at 2-month follow-up, therefore: **1) Capnometry-assisted breathing training + in vivo exposure** (classified as BT, n = 24) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 9

Meuret 2010 *(Continued)*

- **Duration of intervention:** 16 weeks

2) Cognitive training + in vivo exposure (classified as CBT, n = 23)

- **Therapy format:** individual
- **Duration of each session:** 60 minutes
- **Mean number of sessions:** 9
- **Duration of intervention:** 16 weeks

Outcomes	**Time points for assessment:** baseline, pre-phase 1, post-phase 1/pre-exposure (beginning of phase 2), post-exposure (after the 3 weekly sessions of in vivo exposure), 2 months follow-up (last exposure session) **Measures:** Panic Disorder Severity Scale (PDSS), Body Sensations Questionnaire (BSQ), Anxiety Sensitivity Index (ASI), combined measure of symptom appraisal (ASI/BSQ), Mobility Inventory for Agoraphobia (MI), Anxiety Control Questionnaire (ACQ), end-tidal pCO_2, respiration rate (RR) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers at termination (last exposure session) **Continuous scale:** although measured, detailed data are not reported **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"Randomization software was used to assign patients to condition."*
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	In the CT + exposure arm, a total of 12 patients (out of 23 randomised) completed all sessions; in the CART + exposure arm, a total of 16 patients (out of 24 randomised) completed all sessions.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"CART is based on the theory that sustained levels of hypocapnia contribute to symptom development and maintenance of PD (* Meuret 2008 *)."*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All sessions were audio- or video-taped and discussed in the weekly supervision meetings by expert clinicians to ensure that therapists adhered to the treatment protocol. A random sample of 10% of all recorded treatment sessions (10 CART and 10 CT sessions) was evaluated blindly for protocol adherence by two independent, experienced master's-level clinicians. In addition, 50% of the rated sessions (5 CART and 5 CT) were randomly selected and rated by another master's-level and one doctoral-level clinician to assess interrater reliability. Adherence to the given model/protocol was rated high for both conditions (CART: M*

Meuret 2010 *(Continued)*

6.25, SD 0.92; CT: M 5.50, SD 1.20), and ratings were not significantly different between conditions. Interrater agreement was calculated with intraclass correlation coefficients. The results suggest that coders showed high agreement (intraclass correlation [2, 1] 0.85; Shrout & Fleiss, 1979)."

Meyerbroker 2011

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 65 years, DSM-IV-TR diagnosis of panic disorder with agoraphobia **Exclusion criteria:** presence of a medical condition (pregnancy, seizure disorder, pacemaker), current use of tranquillisers or not on a stable dose of antidepressants, current substance dependence, current depression with suicidal ideation, bipolar disorder, borderline or antisocial personality disorder, history of psychosis, severe cognitive impairment **Characteristic of the sample:** • **Age:** not specified • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 55) were randomly assigned to either: **1) CBT + virtual reality exposure** (classified as CBT, n = 19) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** not specified **2) CBT + in vivo exposure** (classified as CBT, n = 18) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 10 • **Duration of intervention:** not specified **3) Wait list** (classified as WL, n = 18) • **Mean number of sessions:** 0 • **Duration of intervention:** 8 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment. **Measures:** Panic Disorder Severity Scale (PDSS), Mobility Inventory (MI), Bodily Sensation Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), Panic Appraisal Inventory (PAI), Avoidance Scale of Watson and Marks The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers

Meyerbroker 2011 *(Continued)*

	Continuous scale: although measured, data cannot be used (n, mean and SD are reported only for the comparison VRET versus IVExp, both classified as CBT: comparison not feasible) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"Patients were randomized with randomly permuted blocks"* (personal communication)
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	High risk	*"After randomization, 9 of the 55 patients declined to start the treatment. During treatment 15 patients dropped out for various reasons: medical issues (n = 2), acute crisis 2), could not experience the virtual environment as real (n = 2), had no further complaints (n = 3) and personal circumstances (n = 1). Completer analyses included the remaining sample of 31 patients who completed all treatment sessions (n = 29) or at least 80% of the sessions (n = 2). Intent-to-treat analyses were done with multiple imputation on the full sample (n = 46) who started at least 1 treatment session"*
Selective reporting (reporting bias)	High risk	Study protocol available: declared primary outcomes are PDSS and MI (plus Behavioural Avoidance Test, added subsequently). Primary outcomes are not specified in published report.
Researcher allegiance	Low risk	*"None of the authors has any conflict of interest"*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Milrod 2006a

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** diagnosis with primary DSM-IV panic disorder with or without agoraphobia, minimum severity score of 5 on the 0- to 8-point Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV-Lifetime Version), minimum of 1 weekly panic attack. Patients were included whether or not they were taking anti-panic medication: subjects meeting study entrance criteria while taking stable doses of medication agreed to keep medication type and dose constant throughout the study. Patients discontinued ongoing psychotherapy to gain study entrance. Patients with comorbid major depression, personality disorders and severe agoraphobia were included. **Exclusion criteria:** psychosis, bipolar disorder and active substance abuse (6 months remission necessary for inclusion) **Characteristic of the sample:** • **Age by treatment group:** PFPP group mean age 33.4 (SD 9.6); ART group mean age 33.5 (SD 8.5) • **Percentage of agoraphobic patients:** PFPP group 69%; ART group 86% • **Percentage of patients on drug therapy:** PFPP group 19%; ART group 17%

Milrod 2006a *(Continued)*

- **Percentage of patients with major depression:** PFPP group 19%; ART group 26%

Interventions	Participants (n = 49) were randomly assigned to either: **1) Panic focused psychodynamic psychotherapy** (classified as PD, n = 26) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 24 • **Duration of intervention:** 12 weeks **2) Applied relaxation training** (classified as PT, n = 23) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 24 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, termination, 2 months follow-up, 4 months follow-up, 6 months follow-up, 12 months follow-up **Measures:** Panic Disorder Severity Scale (PDSS), Sheehan Disability Scale (SDS), Hamilton Depression Rating Scale (HAM-D), Hamilton Anxiety Rating Scale (HAM-A) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** 40% reduction in PDSS at termination **ST-Dropouts:** non-completers **Continuous scale:** Panic Disorder Severity Scale (PDSS) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"Subjects were randomly assigned using a computer generated treatment assignment list that was stratified by presence or absence of 1) comorbid current DSM-IV major depression and 2) stable doses of antipanic medication."*
Allocation concealment (selection bias)	Low risk	*"All study staff was blinded"* (personal communication)
Blinding of outcome assessment (detection bias) ST-Remission	Low risk	*"Independent evaluators, blinded to subject condition and therapist orientation, assessed subjects at baseline, treatment termination, and at 2, 4, 6, and 12 months posttreatment termination"*
Incomplete outcome data (attrition bias) Short-term	High risk	*"Rates of dropout from the 12-week randomized controlled clinical trial differed significantly between the randomly assigned treatment groups: two out of 26 (7%) panic-focused psychodynamic psychotherapy subjects and eight out of 23 (34%) applied relaxation training subjects dropped out. The analyses described above adhered to the intention-to-treat principle using last observation forward to impute missing data for the primary outcome and three continuous secondary*

Milrod 2006a *(Continued)*

		outcomes." Although a LOCF method was used, LOCF cases are many and unbalanced between arms.
Selective reporting (reporting bias)	Low risk	Study protocol available. All of the study's pre-specified (primary and secondary) outcomes have been reported in the pre-specified way.
Researcher allegiance	High risk	Milrod B and Busch F are co-authors of PFPP manual (see Milrod 1997)
Therapist allegiance	Unclear risk	Insufficient information provided
Treatment fidelity	Unclear risk	*"Therapists in both modalities met monthly for group supervision and received individual supervision as needed. Therapists in both modalities were monitored for adherence to treatment protocol by adherence raters in each modality with equal frequency. Three videotapes were rated for adherence per individual treatment. All therapists met predetermined adherence standards."* It is unclear whether all sessions for both therapies were recorded and if the 3 videotapes monitored for adherence where chosen randomly.

Muncy 1991

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with or without agoraphobia, drug-free, in reasonably good health, not currently in therapy (or, in this case, able to arrange with the therapist to work within the structure of the project) **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** not specified • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified (68.4% imputed from BDI in the 3 arms considered for this review)
Interventions	Participants (n = 26) were randomly assigned to either: **1) Rational emotive therapy** (classified as CT, n = 7) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks **2) Rational emotive therapy + biofeedback** (classified as CT, n = 8) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks **3) Imipramine** (not included, n = 7) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** not specified

Muncy 1991 *(Continued)*

- **Duration of intervention:** 6 weeks

4) Wait list (classified as WL, n = 4)

- **Mean number of sessions:** 0
- **Duration of intervention:** 6 weeks

Outcomes	**Time points for assessment:** baseline, mid-treatment (3 weeks after commencement), termination (6 weeks after commencement). **Measures:** Minnesota Multiphasic Personality Inventory (MMPI), Millon Clinical Multiaxial Inventory (MCMI), Beck Depression Inventory (BDI), State-Trait Anxiety Inventory (STAI) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers **Continuous scale:** not measured (none of the available measures meets our inclusion criteria, see Secondary outcomes) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Low risk	*"Two subjects from G3* (Imipramine arm)*abandoned the study in the first three weeks of the treatment due to their inability to tolerate the side-effects of imipramine. Thus, the number of observations was reduced to 24 [...]."* Since we are not considering the imipramine arm for this review, the number of dropouts is 0
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Ost 1993

Methods	**Study design:** randomised controlled trial

Ost 1993 *(Continued)*

Participants	**Inclusion criteria:** age between 20 and 60 years, DSM-III-R diagnosis of panic disorder with agoraphobia, express a willingness to participate in the study for a period of 3 months. If any medication was used the intake was to be held constant during the study; participants had to agree not to receive any other kind of psychiatric or psychological treatment except for any ongoing medication during the treatment. **Exclusion criteria:** any other psychiatric complaint in need of immediate treatment; psychotic or organic symptoms **Characteristic of the sample:** • **Age:** not reported for ITT sample (among completers, mean age was 37.42 years, SD 8.38, range 23 to 56) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not reported for ITT sample (51% among completers) • **Percentage of patients with major depression:** not reported for ITT sample (38% among completers)
Interventions	Participants (n = 46) were randomly assigned to either: **1) Applied relaxation** (classified as PT, n = 15) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) Exposure in vivo** (classified as BT, n = 16) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **3) Cognitive treatment** (classified as CBT, n = 15) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 1 year follow-up **Measures:** Agoraphobia Scale, Fear Questionnaire (FQ), Mobility Inventory for Agoraphobia (MI), Agoraphobic Cognitions Questionnaire (ACQ), Body Sensations Questionnaire (BSQ), Hamilton Anxiety Scale (HAM-A), State-Trait Anxiety Inventory (STAI), Hamilton Depression Scale (HAM-D), Beck Depression Inventory (BDI), Behavior Test (BAT), Autonomic Perception Questionnaire (APQ) The following outcomes were used for quantitative analyses: **ST-Remission:** clinically significant improvement* (Agoraphobia Scale criteria) at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ) **LT-Remission/Response:** clinically significant improvement* (Agoraphobia Scale criteria) at 1 year follow-up

Ost 1993 *(Continued)*

Notes	* *"To assess the degree of clinically significant improvement achieved by the patients the method described by Jacobson, Follette and Revenstorf (1984) was used. On any single measure a patient's post-treatment or follow-up score must be outside the range of the patient group's pre-treatment scores, or inside a normal group's range, in the direction of functionality, defined as Mean ±1.96 x SD . Besides, the change must be statistically reliable. For this study we chose to apply two criteria; percentage of situations completed in the BAT, and score on the avoidance part of Agoraphobic Scale. The respective cutoff scores were:* *BAT: Mean 23.23, SD 14.57, criterion: 22.23 + 1.96(14.57) = 51.79, i.e. 52.* *Agoraphobia Scale-Avoidance: Mean 20.56, SD 6.76, criterion 20.56 - 1.96(6.76) = 7.31."*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	The Agoraphobia Scale (used to determine ST-Remission) is a self report measure
Incomplete outcome data (attrition bias) Short-term	Low risk	*"One patient from the Exposure group dropped out after 2 sessions and was replaced. All the AR- and CT-patients completed the study. At follow-up I patient (in the AR-group) had died, I refused participation in the assessment (E-group), and 1 (AR-group) had moved and was unreachable by mail or telephone. Thus, follow-up assessment was done on 42 (93.3%) of the original patients."*
Incomplete outcome data (attrition bias) Long-term	Low risk	(See above). Missing outcome data balanced in numbers across intervention groups; the proportion of missing outcomes compared with the observed event risk is not enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Ost LG is involved in conceptualisation of Applied Relaxation (see Ost 1987)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided
Other bias	Unclear risk	Modification of the original sample with one replacement

Ost 1995

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 60 years, DSM-III-R diagnosis of panic disorder with mild or no agoraphobic avoidance, duration of the panic disorder of at least 1 year, at least 3 panic attacks during the 3 weeks baseline, panic disorder seen as the patient's primary problem, agreeing to take part in the study for 18 weeks, including pre- and post-assessment, and 1 year follow-up, and be willing to

Ost 1995 *(Continued)*

	accept random allocation. If on prescribed drugs for panic disorder: (a) the dosage had to be constant for 3 months before the start of the treatment; and (b) the patient had to agree to keep the dosage constant throughout the study. **Exclusion criteria:** primary depression (i.e. onset before the start of the panic disorder), any other psychiatric disorder in immediate need of treatment **Characteristic of the sample:** • **Age:** not reported for ITT sample (among completers, mean age was 32.6 years, SD 7.1, range 23 to 45) • **Percentage of agoraphobic patients:** 21.05% • **Percentage of patients on drug therapy:** not reported for ITT sample (72% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 38) were randomly assigned to either: **1) Applied relaxation** (classified as PT, n = 19) • **Therapy format:** individual • **Duration of each session:** 55 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) Cognitive therapy** (classified as CBT, n = 19) • **Therapy format:** individual • **Duration of each session:** 55 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 1 year follow-up. **Measures:** Anxiety Disorders Interview Schedule-Revised (ADIS-R), Hamilton Anxiety Scale (HAM-A), Hamilton Depression Scale (HAM-D), panic diary, Panic Attack Scale (PAS), Beck Anxiety Inventory (BAI), State-Trait Anxiety Inventory (STAI), Self-Rating of Anxiety Scale (SAS), Beck Depression Inventory (BDI), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), Bodily Sensations Interpretations Questionnaire The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Body Sensations Questionnaire (BSQ) **LT-Remission/Response:** high end-state functioning* at 1 year follow-up
Notes	* *"High end-state functioning (HEF) was defined as being panic-free and having an independent assessor rating of severity of the panic disorder of ≤2 (i.e. 'slight') on the* (ADIS-R)*0-8 scale."*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information reported

Ost 1995 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No information reported
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	*"High end-state functioning (HEF) was defined as being panic-free and having an independent assessor rating of severity of the panic disorder of ≤2 (i.e. 'slight') on the* (ADIS-R)*0-8 scale."* It is unclear whether the assessor was blind to treatment allocation.
Incomplete outcome data (attrition bias) Short-term	Low risk	2 patients, both in the AR group, dropped out at an early stage of treatment due to scheduling difficulties. All 36 patients that completed the study were followed up 1 year later. The proportion of missing outcomes compared with the observed event risk is not enough to have a clinically relevant impact on the intervention effect estimate.
Incomplete outcome data (attrition bias) Long-term	Low risk	(see above)
Selective reporting (reporting bias)	High risk	Study protocol unavailable. Agoraphobic Cognitions Questionnaire (ACQ) results are not reported, although this should be an administered scale according to the methods section.
Researcher allegiance	High risk	Ost LG is involved in conceptualisation of Applied Relaxation (see Ost 1987)
Therapist allegiance	Unclear risk	No information reported
Treatment fidelity	Unclear risk	No information reported

Ost 2004

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 60 years, primary DSM-IV diagnosis of panic disorder with agoraphobia, severity of at least 4 on the ADIS-IV 0 to 8 scale, minimum of 1 year duration of the phobia. If on psychotropic medication: have been on a constant dose for at least 4 months and accept to keep the dosage constant throughout therapy. **Exclusion criteria:** primary major depression (i.e. onset before the PDA), current alcohol or substance abuse, psychotic or organic symptoms, ongoing psychotherapy **Characteristic of the sample:** • **Age:** mean 36.1 years (SD 10.3, range 18 to 58) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** 52% • **Percentage of patients with major depression:** 12.3%
Interventions	Participants (n = 73) were randomly assigned to either: **1) Exposure in vivo** (classified as BT, n = 25; after re-randomisation of WL n = 35) • **Therapy format:** individual • **Duration of each session:** 67.5 minutes • **Mean number of sessions:** 14 (12 to 16) • **Duration of intervention:** 14 (12 to 16) weeks **2) Cognitive behaviour therapy** (classified as CBT, n = 26; after re-randomisation of WL n = 36)

Ost 2004 *(Continued)*

- **Therapy format:** individual
- **Duration of each session:** 67.5 minutes
- **Mean number of sessions:** 14 (12 to 16)
- **Duration of intervention:** 14 (12 to 16) weeks

3) Wait list (classified as WL, n = 22)

- **Mean number of sessions:** 0
- **Duration of intervention:** 16

Outcomes	**Time points for assessment:** baseline, post-treatment, 1 year follow-up. **Measures:** Anxiety Disorders Interview Schedule-IV (ADIS-IV), Hamilton Anxiety Scale (HAM-A), Hamilton Depression Scale (HAM-D), Agoraphobia Scale, Mobility Inventory (MI), Fear Questionnaire (FQ), Panic Attack Scale (PAS), Anxiety Sensitivity Index (ASI), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), Beck Anxiety Inventory (BAI), State-Trait Anxiety Inventory (STAI), Beck Depression Inventory (BDI), Quality of Life Inventory panic diary, behavioural approach tests The following outcomes were used for quantitative analyses: **ST-Remission:** no longer meet DSM-IV criteria for PDA at post-treatment (original sample) **ST- Response:** not measured (imputed from the continuous scale for the original sample) **ST-Dropouts:** non-completers (original sample) **Continuous scale:** Anxiety Sensitivity Index (ASI) measured on the original sample (SDs are not reported but could be borrowed from other studies using ASI) **LT-Remission/Response:** no longer meet DSM-IV criteria for PDA at 1 year follow-up (original + re-randomised sample)
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	Unclear risk	*"An independent research assistant not involved with the treatment performed all the diagnostic interviews and ratings."* It is unclear whether the assessor was blind to treatment allocation.
Incomplete outcome data (attrition bias) Short-term	Low risk	*"Seven patients dropped out of treatment before completion; 3 (13%) in the E-group, 2 (8%) in the CBT-group, and 2 (9%) in the WLC-group, a non-significant difference."* Missing outcome data balanced in numbers across intervention groups. The proportion of missing outcomes compared with the observed event risk is not enough to have a clinically relevant impact on the intervention effect estimate.
Incomplete outcome data (attrition bias) Long-term	High risk	*"After the post-assessment of the waiting list two patients dropped out and the remaining were randomized to the two treatment conditions with 10 in each. When the former WLC-patients were treated 3 of the 10 randomized to the E-group, but none in the CBT-group, dropped out."* Reason for missing outcome

Ost 2004 *(Continued)*

		data likely to be related to true outcome, with imbalance in numbers for missing data across intervention groups.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"The therapy sessions were audiotaped and upon completion, three sessions were randomly selected for every patient, one from each third of the therapy, for analysis. Three licensed psychotherapists with long experience of CBT and blind to the allocation of the patients listened to the tapes and classified the sessions as either an exposure or a CBT-session. They also rated the competency of the therapists using a modified version of the Cognitive therapy scale (Young & Beck, 1980), consisting of nine specific items and a general item (rated on a 0–6 scale). Which treatment the patient was receiving was correctly identified in 96.7% of the sessions rated by the experts, indicating a satisfying degree of treatment integrity."*

Petterson 1996

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** range 20 to 63 years • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 27*) were randomly assigned to either: **1) Cognitive behaviour treatment** (classified as CBT, n = 14?) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 6 • **Duration of intervention:** 6 weeks **2) No treatment** (classified as NT, n = 13?) • **Mean number of sessions:** 0 • **Duration of intervention:** 6 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** Anxiety Sensitivity Index (ASI), State-Trait Anxiety Inventory (STAI), Panic Attack Record, physiologic measures (blood pressure, pulse rate, finger temperature) The following outcomes were used for quantitative analyses:

Petterson 1996 *(Continued)*

	ST-Remission: not measured (not imputed because number of randomised patients is unclear) **ST- Response:** not measured (not imputed because number of randomised patients is unclear) **ST-Dropouts:** unclear **Continuous scale:** Anxiety Sensitivity Index (ASI). SDs are not reported but could be borrowed from other studies using ASI **LT-Remission/Response:** not measured
Notes	*It is unclear whether the reported number participants (n = 27) refers to the ITT sample or to treatment completers (*"Twenty-seven adults completed the study. [...] The subjects were randomly assigned to either the Treatment (n = 14) or Control (n = 13) conditions"*).

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Unclear risk	No information provided
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Rees 1999

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III diagnosis of panic disorder with little or no agoraphobic avoidance **Exclusion criteria:** secondary diagnosis with an overall severity rating less than 2 points away from the panic disorder severity rating on the clinician's 9-point rating scale in ADIS; any medical condition such as asthma, angina, emphysema that might have complicated the panic disorder **Characteristic of the sample:** • **Age:** not specified • **Percentage of agoraphobic patients:** 82.5% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** 11.1%
Interventions	Participants (n = 40) were randomly assigned to either:

Rees 1999 *(Continued)*

	1) Information giving + self monitoring (classified as PE, n = 20) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 2 • **Duration of intervention:** 2 weeks **2) Self monitoring alone** (classified as APP, n = 20) • **Therapy format:** not specified • **Duration of each session:** 60 minutes • **Mean number of sessions:** 2 • **Duration of intervention:** 2 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment. **Measures:** daily records of panic, anxiety, depression and anticipatory fear of panic The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (not imputed because of skewed distribution of available continuous scale) **ST- Response:** not measured (not imputed because of skewed distribution of available continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** panic frequency **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Low risk	No dropouts
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Reinecke 2013

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-IV diagnosis of panic disorder with or without agoraphobia, naive to exposure-based CBT; at least moderate agoraphobic avoidance, panic-related safety behaviours (e.g. medication, or standing close to an escape exit to prevent panic attacks) and catastrophic panic cognitions (e.g. "If I stay here my heart will beat even faster, and I will suffer a heart attack"), assessed with a structured panic assessment interview. Occasional medication with benzodiazepines or β-blockers as needed was not an exclusion criterion; however, patients were required to be medication-free 48 hours before the test sessions to avoid any interference with experimental testing and CBT. **Exclusion criteria:** lifetime history of epilepsy, psychotic disorders, bipolar disorder or substance abuse, primary depressive disorder, insufficient English skills, psychopharmacological or psychotherapeutic treatment during the last 6 months **Characteristic of the sample:** • **Age by group:** treatment group mean age 36 (SD 14.7); wait list group mean age 35.1 (SD 14.1) • **Percentage of agoraphobic patients:** treatment group 100%; wait list group 85.7% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** 7.14%
Interventions	Participants (n = 28) were randomly assigned to either: **1) Single session exposure-based CBT** (classified as CBT, n = 14) • **Therapy format:** not specified • **Duration of each session:** 30 minutes • **Mean number of sessions:** 1 • **Duration of intervention:** not applicable (single session) **2) Wait list** (classified as WL, n = 14) • **Mean number of sessions:** 0 • **Duration of intervention:** 4 weeks
Outcomes	**Time points for assessment:** baseline (day 1), post-treatment (day 2), follow-up (4 weeks) **Measures:** Hospital Anxiety and Depression Scale, Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ), Mobility Inventory (MI), Faces Dot Probe Task, Stress Test The following outcomes were used for quantitative analyses: **ST-Remission:** MI-Ag score falling within the range reported for healthy control subjects **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ) **LT-Remission/Response:** not measured
Notes	None
Risk of bias	

Bias	Authors' judgement	Support for judgement

Reinecke 2013 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Mobility Inventory (used to determine ST-Remission) is a self rated measure.
Incomplete outcome data (attrition bias) Short-term	Low risk	All patients were assessed at 4 weeks follow-up.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"Our single-session treatment was a very condensed version of psychological intervention recommended for delivery in routine clinical care"*
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Salkovskis 1999

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with moderate or severe avoidance, at least 2 panic attacks occurring in the 4 weeks prior to assessment, a score of 9 or more on the modified Fear Questionnaire, agoraphobic avoidance sub-scale (FQ-Ag), being unable to complete the penultimate step of a pre-determined standardised behavioural avoidance test course conducted prior to the experimental procedure, it was possible to identify both catastrophic thoughts which occurred during panic attacks, and safety-seeking behaviours which the patients said they carried out during the attacks to prevent the feared catastrophes. In addition, it was required that the patient rated an increase of anxiety from baseline of at least 20 points on a 100 point visual analogue scale when entering the 5-minute individualised behaviour test. **Exclusion criteria:** none **Characteristic of the sample:** • **Age by group:** decreased safety behaviours group mean age 42.11 (SD 13.5); exposure only group mean age 33.6 (SD 11.7) • **Percentage of agoraphobic patients:** 100% • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified (44% according to imputation from BDI)
Interventions	Participants (n = 18) were randomly assigned to either: **1) Cognitive behaviour therapy** (classified as CBT, n = 9) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 4

Salkovskis 1999 *(Continued)*

- **Duration of intervention:** 8 days

2) Habituation-based exposure therapy (classified as BT, n = 9)

- **Therapy format:** not specified
- **Duration of each session:** not specified
- **Mean number of sessions:** 4
- **Duration of intervention:** 8 days

Outcomes	**Time points for assessment:** baseline, post-treatment. **Measures:** Beck Depression Inventory (BDI), Beck Anxiety Inventory (BAI), panic frequency, Fear Questionnaire (FQ), Agoraphobic Cognitions Questionnaire (ACQ), standardised behavioural walk (BW), individualised behavioural test (BT). The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"Randomisation was on the basis of sampling without replacement, using sealed envelopes opened on completion of the initial assessments"*
Allocation concealment (selection bias)	Unclear risk	It is unclear whether envelopes were opaque and sequentially numbered.
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Agoraphobic Cognitions Questionnaire (ACQ), used to impute ST-remission, is a self rated measure.
Incomplete outcome data (attrition bias) Short-term	Unclear risk	Missing outcome data balanced in numbers across intervention groups (1 dropout from each arm), it is unclear whether the proportion of missing outcomes compared with the observed event risk is enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	Both Salkovskis PM and Clark DM are authors of a manual for CBT in panic disorder (see Clark 1986b)
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Schmidt 1997a

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** principal DSM-IV diagnosis of panic disorder **Exclusion criteria:** change in medication type or dose during the 8 weeks preceding entry, evidence of serious suicide intent, current substance abuse, current or past schizophrenia, bipolar disorder, organic mental disorder **Characteristic of the sample:** • **Age:** not specified • **Percentage of agoraphobic patients:** 58% • **Percentage of patients on drug therapy:** 59% • **Percentage of patients with major depression:** not specified (16.6% according to imputation from BDI)
Interventions	Participants (n = 54) were randomly assigned to either: **1) Cognitive behaviour therapy + respiratory training** (classified as CBT, n = 18) • **Therapy format:** group • **Duration of each session:** not specified • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **2) Cognitive behaviour therapy** (classified as CBT, n = 20) • **Therapy format:** group • **Duration of each session:** not specified • **Mean number of sessions:** 12 • **Duration of intervention:** 12 weeks **3) Wait list** (classified as WL, n = 16) • **Mean number of sessions:** 0 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment (12 weeks) **Measures:** Clinical Global Impression - Severity (CGI-S), panic diary, Shehan Patient-Rated Scale (SPRAS), Fear Questionnaire (FQ), Anxiety Sensitivity Inde (ASI), Panic Appraisal Inventory (PAI), Beck Depression Inventory (BDI), Shehan Disability Scale (SDS), Acute Panic Inventory (API), physiological measures (vital capacity, CO_2 intake volume), psychophysiological measures (heart rate, systolic blood pressure, diastolic blood pressure) The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Inde (ASI) **LT-Remission/Response:** not measured
Notes	* High end-state functioning: *"A patient was classified as recovered when scores on each of three clinical dimensions (i.e. panic frequency, anxiety and phobic avoidance) fell within the normal range. Recovery criteria for the SPRAS and FQ are based on established cutoff scores reported in the literature. The recovery criterion for panic attacks was zero."*

Schmidt 1997a *(Continued)*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	The 3 measures used to assess the high end-state functioning status (panic frequency, SPRAS score, FQ score) are all self reported.
Incomplete outcome data (attrition bias) Short-term	High risk	*"Of those patients assigned to the treatment conditions* (n=54)*, 34 were assessed at post-treatment."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"Treatment integrity was maintained by using a structured and manualized treatment protocol (* Schmidt 1994 *) that described specific goals and strategies for each session."* Schmidt NB is among the study authors.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	High risk	*"Treatment integrity was maintained by using a structured and manualized treatment protocol that described specific goals and strategies for each session."*

Schmidt 1997b

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** principal DSM-IV diagnosis of panic disorder with or without agoraphobia (active duty military sample) **Exclusion criteria:** change in medication type or dose during the 8 weeks preceding entry, evidence of serious suicide intent, current substance abuse, current or past schizophrenia, bipolar disorder, organic mental disorder **Characteristic of the sample:** • **Age:** mean 31.8 years (SD 9.7) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified (20% according to imputation based on BDI)
Interventions	Participants (n = 37) were randomly assigned to either: **1) Cognitive behaviour treatment** (classified as CBT, n = 25) • **Therapy format:** group • **Duration of each session:** not specified

Schmidt 1997b *(Continued)*

- **Mean number of sessions:** 12
- **Duration of intervention:** 12 weeks

2) Wait list (classified as WL, n = 12)

- **Mean number of sessions:** 0
- **Duration of intervention:** 12 weeks

Outcomes	**Time points for assessment:** baseline (week 0), post-treatment (week 9), 3 months follow-up (week 21) **Measures:** Texas Panic Attack Record Form, Shehan Patient-Rated Scale (SPRAS), Fear Questionnaire (FQ), Shehan Disability Scale (SDS), Beck Depression Inventory (BDI) The following outcomes were used for quantitative analyses: **ST-Remission:** high end-state functioning* at post-treatment **ST- Response:** not measured **ST-Dropouts:** non-completers **Continuous scale:** not measured **LT-Remission/Response:** not measured
Notes	* High end-state functioning: *"A patient was classified as recovered when scores on each of three clinical dimensions (i.e. panic frequency, anxiety and phobic avoidance) fell within the normal range. Recovery criteria for the SPRAS and FQ are based on established cutoff scores reported in the literature. The recovery criterion for panic attacks was set at zero."*

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	The 3 measures used to assess the high end-state functioning status (panic frequency, SPRAS score, FQ score) are all self reported.
Incomplete outcome data (attrition bias) Short-term	High risk	*"Of those patients assigned to a treatment condition (n=37), 29 were assessed at post-treatment. Dropouts were similar across conditions, with 20% of patients in the immediate treatment condition (n=5) and 25% of patients in the delayed treatment condition (n=3) discontinuing their participation."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"Treatment integrity was maintained by utilizing a structured and manualized treatment protocol (* Schmidt 1994 *) that described specific goals and strategies for each session."* Schmidt NB is among the study authors.
Therapist allegiance	Unclear risk	No information provided

Schmidt 1997b *(Continued)*

Treatment fidelity	High risk	*"Treatment integrity was maintained by utilizing a structured and manualized treatment protocol that described specific goals and strategies for each session."*

Scott 1995

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of generalised anxiety disorder, panic disorder with or without agoraphobia and major depression **Exclusion criteria:** none **Characteristic of the sample (panic disorder sub-sample):** • **Age:** not specified • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified
Interventions	Participants (panic disorder sub-sample, n = 21) were randomly assigned to either: **1) Cognitive behaviour therapy** (classified as CBT, n = 15) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 7 • **Duration of intervention:** 7 weeks **2) Wait list** (classified as WL, n = 6) • **Mean number of sessions:** 0 • **Duration of intervention:** 7 weeks
Outcomes	**Time points for assessment:** baseline, start of treatment, termination, 6 weeks follow-up, 3 months follow-up, 6 months follow-up, 12 months follow-up **Measures:** Beck Anxiety Inventory (BAI), Beck Depression Inventory (BDI) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers **Continuous scale:** not measured **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement

Scott 1995 *(Continued)*

Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Incomplete outcome data (attrition bias) Short-term	Low risk	No dropouts
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Sharp 2004

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 70 years, DSM-IV diagnosis of panic disorder with or without agoraphobia, score a minimum of 15 on the Hamilton Anxiety Scale, score a maximum of 20 on the Montgomery Asberg Depression Rating Scale, symptoms lasting at least 3 months. Patients taking concurrent psychotropic medications were not excluded from the study, but were required to continue taking these medications as prescribed throughout the study period. **Exclusion criteria:** having received a psychological treatment for panic disorder and agoraphobia in the past 6 months **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age ranged from 34.6 to 41.7 depending on group) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified for ITT sample (50% among completers) • **Percentage of patients with major depression:** not specified
Interventions	Participants (n = 97) were randomly assigned to either: **1) Group CBT** (classified as CBT, n = 38) • **Therapy format:** group • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 12 weeks **2) Individual CBT** (classified as CBT, n = 37) • **Therapy format:** group • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 12 weeks

Sharp 2004 *(Continued)*

	3) Wait list (classified as WL, n = 22) • **Mean number of sessions:** 0 • **Duration of intervention:** 12 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 3 months follow-up **Measures:** Hamilton Anxiety Scale (HAM-A), Symptom Rating Test (SRT), Montgomery Asberg Depression Rating Scale (MADRS), Fear Questionnaire - agoraphobia sub-scale (FQ-Ag), panic diary The following outcomes were used for quantitative analyses: **ST-Remission:** FQ-Ag below 10 at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Fear Questionnaire - agoraphobia sub-scale (FQ-Ag) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Although assessor was blind to treatment allocation, FQ (used to extract ST-Remission), is a self rated measure
Incomplete outcome data (attrition bias) Short-term	High risk	*"A total of n = 27 patients dropped out of treatment early failing to complete five sessions of study treatment. Group treatment had a significantly higher drop-out rate (n=18, 47%), than either individual treatment (n=6, 16%), or waiting list (n=3, 14%)."*
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Unclear risk	*"For the two treatment groups all patients received the same CBT and all received identical treatment instructions with the same written treatment manual being supplied to all patients."* It is unclear whether authors were involved in writing the administered manual.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Shear 1994

Methods	**Study design:** randomised controlled trial

Shear 1994 *(Continued)*

Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with or without agoraphobia. All subjects were required to discontinue psychotropic medication for at least 2 weeks before study entry and to refrain from using medication or any other psychotherapeutic treatment during the study. **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** not specified for ITT sample (among completers, mean age was 34.7 years, SD 9.7) • **Percentage of agoraphobic patients:** 92% • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified (18.2% according to imputation from BDI)
Interventions	Participants (n = 66) were randomly assigned to either: **1) Cognitive behaviour treatment** (classified as CBT, n = 37) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks **2) Non-prescriptive treatment** (classified as SP, n = 29) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 15 • **Duration of intervention:** 15 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment, 6 months follow-up **Measures:** Anxiety Disorders Interview Schedule–R (ADIS-R), panic diary, Brief Fear Questionnaire, Mobility Inventory (MI), Beck Depression Inventory (BDI), Hopkin's Symptom Checklist (SCL-90), Sheehan Disability Scale (SDS), Hamilton Rating Scale for Anxiety (HAM-A), Hamilton Rating Scale for Depression (HAM-D), State-Trait Anxiety Inventory (STAI), Anxiety Sensitivity Index (ASI), Body Sensations Questionnaire (BSQ), Agoraphobic Cognitions Questionnaire (ACQ) The following outcomes were used for quantitative analyses: **ST-Remission:** panic-free at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** although measured (panic-free at 6 months follow-up) data were not entered in the analyses because dropouts exceeded 30% of the originally randomised sample (see Secondary outcomes).
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided

Shear 1994 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	*"Most pretreatment and all posttreatment and follow-up interviews were conducted by an independent evaluator who was "blind" to the treatment condition, study aims, and methods."* However, ST-remission is defined as being panic-free, a self rated measure.
Incomplete outcome data (attrition bias) Short-term	High risk	*"Pretreatment data were obtained on 45 subjects (CBT=24, NPT=21). Forty-one subjects completed posttreatment assessments (CBT=20, NPT=21)."*
Incomplete outcome data (attrition bias) Long-term	High risk	*"43 subjects completed the follow-up assessments (CBT=23, NPT=20)."*
Selective reporting (reporting bias)	High risk	Study protocol unavailable. The results of some measures planned in the methods section are not reported (e.g. ACQ, BSQ).
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All treatments were supervised by one of us (MKS) in weekly meetings. The CBT sessions were rated for treatment adherence by Michelle Craske, PhD, and Karla Moras, PhD, from the Albany Stress and Anxiety Clinic. The NPT sessions were rated for absence of CBT material. The CBT intervention was rated separately for protocol adherence, general skill, and specific skill on a scale of 0 through 8. The NPT intervention was screened for absence of CBT. Thirty audiotapes were randomly selected from each condition. The mean adherence ratings for CBT sessions was 4.5 on a scale of 0 through 8. [...] There was no indication that NPT sessions 3 through 16 included CBT techniques. The NPT intervention was not rated for adherence or skill in this study."*

Taylor 1982

Methods	**Study design:** randomised controlled trial, cross-over design
Participants	**Inclusion criteria:** DSM-III diagnosis of panic disorder, at least 1 panic attack in the last 3 weeks, agreement to discontinue all psychotropic medications except as prescribed in the study **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** mean 34.9 years (range 18 to 25) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** not specified
Interventions	In the first phase of the treatment (before cross-over), participants (n = unclear) were randomly assigned to either: **1) Relaxation** (classified as, n = unclear) • **Therapy format:** not specified

Taylor 1982 *(Continued)*

- **Duration of each session:** 30 minutes
- **Mean number of sessions:** up to 5
- **Duration of intervention:** not specified

2) Diazepam (not included, n = unclear)

- **Therapy format:** not specified
- **Duration of each session:** 30 minutes
- **Mean number of sessions:** 4 to 5
- **Duration of intervention:** not specified

3) Placebo (not included, n = unclear)

- **Therapy format:** not specified
- **Duration of each session:** 30 minutes
- **Mean number of sessions:** 4 to 5
- **Duration of intervention:** not specified

4) No treatment (classified as WL, n = unclear)

- **Mean number of sessions:** 0
- **Duration of intervention:** not specified

Outcomes	**Time points for assessment:** baseline, post-treatment **Measures:** self report diary (anxiety, mood), State-Trait Anxiety Inventory (STAI), Profile of Mood States, physiological measures (heart rate, skin conductance) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** detailed data are not reported **Continuous scale:** the only available measure (STAI) is not considered among the outcomes of interest for this review (see Secondary outcomes). Furthermore, detailed data are not reported. **LT-Remission/Response:** not measured
Notes	

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Unclear risk	It is unclear whether the researchers involved have a vested interest for or against the psychological therapy under study (ref. 1 of study report).
Therapist allegiance	Unclear risk	No information provided

Taylor 1982 *(Continued)*

Treatment fidelity	Unclear risk	No information provided

Telch 1993

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** age between 18 and 65 years, principal DSM-III-R diagnosis of panic disorder with or without agoraphobia, at least 1 panic attack during the past 30 days **Exclusion criteria:** recent change in psychotropic medications; current psychosis, bipolar disorder or substance abuse disorder **Characteristic of the sample:** • **Age:** mean 34.6 years (SD 10.3) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** 61.2% • **Percentage of patients with major depression:** not specified (34.3% according to imputation from BDI)
Interventions	Participants (n = 67) were randomly assigned to either: **1) Panic inoculation training** (classified as CBT, n = 34) • **Therapy format:** group • **Duration of each session:** 90 minutes • **Mean number of sessions:** 12 • **Duration of intervention:** 8 weeks **2) Wait list** (classified as WL, n = 33) • **Mean number of sessions:** 0 • **Duration of intervention:** 8 weeks
Outcomes	**Time points for assessment:** baseline, post-treatment (week 9), 6 months follow-up **Measures:** Sheehan Patient-Rated Anxiety Scale (SPRAS), Agoraphobia scale of the Fear Questionnaire (FQ-Ag), Beck Depression Inventory (BDI), Anxiety Sensitivity Index (ASI), panic diary The following outcomes were used for quantitative analyses: **ST-Remission:** recovery* at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI) **LT-Remission/Response:** not measured
Notes	* *"The two groups were compared on a composite index of recovery defined as the proportion of patients attaining normal levels of functioning on all three major facets of the disorder (i.e. panic attacks, anxiety and panic-related avoidance)."* Recovery criterion for panic attack was set at 0, for anxiety it was SPRAS score < 30, for avoidance it was FQ-Ag < 12.
Risk of bias	

Telch 1993 *(Continued)*

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	The 3 measures used to assess the high end-state functioning status (panic frequency, SPRAS score, FQ score) are all self reported
Incomplete outcome data (attrition bias) Short-term	Low risk	No dropout in either group
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	High risk	*"A 65-page treatment manual (Telch 1990) describes the specific goals and strategies for each session."*
Therapist allegiance	High risk	"All sessions were conducted by one primary therapist (MJT, JL or NBS) and a graduate student assistant." As noted above, MJ Telch and NB Schmidt may have a vested interested in the success of their active treatment.
Treatment fidelity	Low risk	*"To help protect the integrity of the treatment, therapists and their assistants followed a procedural outline for each therapy session. In addition, all treatment sessions were videotaped and randomly selected segments were rated for consistency with the written treatment protocol."*

Tyrer 1988

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III diagnosis of generalised anxiety disorder, panic disorder or dysthymic disorder; currently receiving no psychotropic drugs (and had not been taking benzodiazepines, neuroleptic drugs or antidepressants in regular dosage for at least 4 weeks before entry to the study); willing to enter the study, in which they understood that they would receive either drug or psychological treatments for up to 10 weeks **Exclusion criteria:** major depressive episode or other psychiatric disorders that take diagnostic precedence over the 3 diagnoses above in the DSM-III classification **Characteristic of the panic disorder sub-sample:** • **Age:** not specified (in the full sample median age was 35 years, range 17 to 76) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** 0% • **Percentage of patients with major depression:** 0%
Interventions	Participants (panic disorder sub-sample, n = 74) were randomly assigned to either: **1) Cognitive and behaviour therapy** (classified as CT, n = 33) • **Therapy format:** not specified • **Duration of each session:** 60 minutes

Tyrer 1988 *(Continued)*

- **Mean number of sessions:** 5
- **Duration of intervention:** 6 weeks

2) Self help treatment package (classified as APP, n = 17)

- **Therapy format:** not specified
- **Duration of each session:** 15 minutes
- **Mean number of sessions:** 5
- **Duration of intervention:** 6 weeks

3) Diazepam (not included, n = 7)

- **Therapy format:** not specified
- **Duration of each session:** not specified
- **Mean number of sessions:** not specified
- **Duration of intervention:** 6 weeks

4) Dothiepin (not included, n = 7)

- **Therapy format:** not specified
- **Duration of each session:** not specified
- **Mean number of sessions:** not specified
- **Duration of intervention:** 6 weeks

5) Placebo (not included, n = 10)

- **Therapy format:** not specified
- **Duration of each session:** not specified
- **Mean number of sessions:** not specified
- **Duration of intervention:** 6 weeks

Outcomes	**Time points for assessment:** baseline, mid-treatment (2, 4, 6 weeks), post-treatment (10 weeks), 16 weeks follow-up, 32 weeks follow-up, 1 year follow-up, 2 years follow-up, 5 years follow-up, 12 years follow-up **Measures:** Comprehensive Psychopathological Rating Scale (CPRS), Montgomery & Asberg Depression Rating Scale (MADRS), Brief Anxiety Scale (BAS), Hospital Anxiety and Depression Scale (HADS), General Neurotic Syndrome Scale (GNSS), Life Events Schedule, Personality Assessment Schedule (PAS) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured **ST- Response:** not measured **ST-Dropouts:** non-completers **Continuous scale:** although measured, none of the available scales is considered for this review (see Secondary outcomes) **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	*"The randomised treatments were indicated by opening a sealed envelope once patients satisfied the inclusion criteria for the study."*

Tyrer 1988 *(Continued)*

Allocation concealment (selection bias)	Unclear risk	It is unclear whether envelopes were opaque and sequentially numbered
Incomplete outcome data (attrition bias) Short-term	Low risk	Considering only panic disorder patients allocated to either CBT (n = 33) or self help treatment package (n = 17), 3 patients were lost at post-treatment assessment (personal communication: 2 patients in the CBT arm, 1 patient in the self help arm). Therefore, missing outcome data are low in number and balanced across intervention groups.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Unclear risk	No information provided

Williams 1996

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-III-R diagnosis of panic disorder with or without agoraphobia, at least 2 DSM-III-R-defined panic attacks per week over a 2-week baseline period. Subjects were asked to refrain from receiving any other psychological treatment for panic, phobia or depression during the study until after the 6-week follow-up. Subjects taking regularly prescribed medications were asked not to alter their medication regimen during the programme, and all subjects were asked to refrain from taking discretionary medication. **Exclusion criteria:** none **Characteristic of the sample:** • **Age:** mean 38 years (range 17 to 76) • **Percentage of agoraphobic patients:** 91.67% • **Percentage of patients on drug therapy:** 62.5% • **Percentage of patients with major depression:** 52.08%
Interventions	Participants (n = 48) were randomly assigned to either: **1) Cognitive treatment** (classified as CT, n = 14) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks **2) Performance treatment** (classified as BT, n = 12) • **Therapy format:** individual • **Duration of each session:** 60 minutes • **Mean number of sessions:** 8 • **Duration of intervention:** 8 weeks **3) Combined treatment** (classified as CBT, n = 13)

Williams 1996 *(Continued)*

- **Therapy format:** individual
- **Duration of each session:** 60 minutes
- **Mean number of sessions:** 8
- **Duration of intervention:** 8 weeks

4) Delayed treatment (classified as WL, n = 9)

- **Mean number of sessions:** 0
- **Duration of intervention:** 8 weeks

Outcomes	**Time points for assessment:** baseline, post-treatment, 6 weeks follow-up, 1 to 2 years follow-up **Measures:** panic diary, Self-Efficacy Scales for Agoraphobia (SESA), Fear Questionnaire (FQ), panic coping self efficacy, Agoraphobic Cognitions Questionnaire (ACQ), Body Sensations Questionnaire (BSQ), Beck Depression Inventory (BDI) The following outcomes were used for quantitative analyses: **ST-Remission:** panic-free at post-treatment **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Agoraphobic Cognitions Questionnaire (ACQ) **LT-Remission/Response:** panic-free at 1 to 2 years follow-up
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	Panic frequency is a self rated measure
Incomplete outcome data (attrition bias) Short-term	Low risk	*"All Ss completed the treatment program; there were no dropouts in any treatment condition."*
Incomplete outcome data (attrition bias) Long-term	Unclear risk	Among the 39 subjects allocated to the 3 active intervention arm, 34 were assessed at 1 to 2-year follow-up. Dropouts (n = 5) were 2/14 in cognitive treatment group, 2/12 in the performance treatment group and 1/13 in the combined treatment group. Missing outcome data are balanced across intervention groups. It is unclear whether the proportion of missing outcomes compared with the observed event risk is enough to have a clinically relevant impact on the intervention effect estimate.
Selective reporting (reporting bias)	Unclear risk	Study protocol unavailable

Williams 1996 *(Continued)*

Researcher allegiance	High risk	*"The performance-based treatment was an office-based adaptation of guided mastery treatment (* Williams 1990 *; Williams, Dooseman & Kleifield 1984; Williams & Zane 1989; Zane & Williams 1993), which emphasizes the importance of performance successes in helping people gain a sense of mastery and self-efficacy."* Williams SL is among the study authors.
Therapist allegiance	Unclear risk	Insufficient information provided
Treatment fidelity	Low risk	*"All treatment sessions were audiotaped, and for each treatment condition, one tape from each therapist was chosen randomly, and all therapist statements on it were transcribed and assembled into sets containing five consecutive statements. Two assistants trained in the coding manual independently, coded the sets of statements, while remaining unaware of Ss' assigned treatment condition and of the hypotheses under investigation. Based on the codes of one randomly chosen coder, in the performance treatment sessions, 72% of the sets of therapist statements contained performance interventions and 0% cognitive interventions. Sets of therapist statements in the cognitive treatment sessions contained 58% cognitive interventions and 0% performance interventions. Combined treatment sessions contained 54% performance interventions and 29% cognitive interventions."*

Wollburg 2011

Methods	**Study design:** randomised controlled trial
Participants	**Inclusion criteria:** DSM-IV diagnosis of panic disorder with or without agoraphobia, willing to accept an 8-week treatment delay if assigned to the wait list **Exclusion criteria:** history of schizophrenia, bipolar disorder, dementia, alcohol or drug abuse, current use of medications with pronounced sympathetic, parasympathetic or respiratory effects, current score on the BDI exceeding 30, current suicidality **Characteristic of the sample:** • **Age (by group):** mean 43.8 years (SD10.7); mean 43.7 years (SD 14.5); mean 38.3 years (SD 14.4) • **Percentage of agoraphobic patients:** not specified • **Percentage of patients on drug therapy:** not specified • **Percentage of patients with major depression:** not specified (8.1% imputed from BDI)
Interventions	Participants (n = 74) were randomly assigned to either: **1) Lower-CO_2 breathing retraining** (classified as PT, n = 19) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 5 • **Duration of intervention:** 5 weeks **2) Raise-CO_2 breathing retraining** (classified as PT, n = 28) • **Therapy format:** not specified • **Duration of each session:** not specified • **Mean number of sessions:** 5 • **Duration of intervention:** 5 weeks **3) Wait list** (classified as WL, n = 27)

Wollburg 2011 *(Continued)*

- **Mean number of sessions:** 0
- **Duration of intervention:** 8 weeks

Outcomes	**Time points for assessment:** baseline, 1-month follow-up (coincides with end of wait list), 6-month follow-up **Measures:** Panic Disorder Severity Scale (PDSS), Anxiety Sensitivity Index (ASI), Beck Depression Inventory (BDI), Mobility Inventory for Agoraphobia (MI), Agoraphobic Cognitions Questionnaire (ACQ), Anxiety Control Questionnaire, Beck Anxiety Inventory (BAI), Body Sensations Questionnaire (BSQ) The following outcomes were used for quantitative analyses: **ST-Remission:** not measured (imputed from continuous scale) **ST- Response:** not measured (imputed from continuous scale) **ST-Dropouts:** non-completers **Continuous scale:** Anxiety Sensitivity Index (ASI). Note that reported SDs are uncommonly low, so we considered them as being SEs. **LT-Remission/Response:** not measured
Notes	None

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No information provided
Allocation concealment (selection bias)	Unclear risk	No information provided
Blinding of outcome assessment (detection bias) ST-Remission	High risk	ASI is a self rated measure
Incomplete outcome data (attrition bias) Short-term	Low risk	No dropouts at 1-month follow-up (see study flow chart)
Selective reporting (reporting bias)	High risk	Study protocol available. PDSS scores at 1-month follow-up (primary outcome) are not reported.
Researcher allegiance	Low risk	To our knowledge, the researchers involved do not have a vested interest for or against the psychological therapy under study.
Therapist allegiance	Unclear risk	No information provided
Treatment fidelity	Low risk	*"All sessions were audiotaped. Of the 50 treatment completers, the session for 20 were randomly selected and rated for therapist competence and adherence. Overall competence ratings of all therapists ranged from 3 (good) to 5 (excellent). The mean adherence rating as measured by application of respiratory behavioral techniques was 5.35 (SD 0.67), with rating of 6 being excellent."*

3W: third-wave
AAAS: African American Acculturation Scale - Short Form
AAQ: Acceptance and Action Questionnaire

ACQ: Agoraphobic Cognitions Questionnaire
ADIS–IV: Anxiety Disorders Interview Schedule–IV
API: Acute Panic Inventory
APP: attention/psychological placebo
APQ: Autonomic Perception Questionnaire
AR: applied relaxation
ART: applied relaxation training
ASI: Anxiety Sensitivity Index
ATQ: Attitude Toward Treatment Questionnaire/Automatic Thoughts Questionnaire
AWS: Anxiety and Willingness Scale
BAI: Beck Anxiety Inventory
BATs: Behavioral Avoidance Tests
BBSIQ: Brief Body Sensations Interpretation Questionnaire
BDI: Beck Depression Inventory
BSI: Brief Symptom Inventory
BSQ: Body Sensations Questionnaire
BT: behaviour therapy
CART: capnometry-assisted breathing training
CBT: cognitive behaviour therapy
CCLAS: Anxiety Scale of the Cognitions Checklist
CGI: Clinician Global Impression
CT: cognitive training
DPAR: Daily Panic Attack Records
DSI: Depressive Symptoms Inventory
DSM: Diagnostic and Statistical Manual of Mental Disorders
DSQ: Diagnostic Symptom Questionnaire
ECQ: Emotional Control Questionnaire
EMDR: eye movement desensitisation and reprocessing
FQ: Fear Questionnaire
FQ-Ag: Fear Questionnaire-Agoraphobia Subscale
FSS-IZZ: Fear Surrey Schedule-III
GNSS: General Neurotic Syndrome Scale
HAM-D: Hamilton Depression Scale
HARS: Hamilton Anxiety Rating Scale
HARS-R: Hamilton Anxiety Rating Scale - Revised
HEF: high end-state functioning
HR: heart rate
HRV: heart rate variability
IBT: Irrational Belief Test
ICD-10: International Classification of Diseases
IET: Interoceptive Exposure Test
ITT: intention-to-treat
IVExp: in vivo exposure
LEAS: Levels of Emotional Awareness Scale
LOCF: last observation carried forward
LT: long-term
MAO: monoamine oxidase
MCMI: Millon Clinical Multiaxial Inventory
MMPI: Minnesota Multiphasic Personality Inventory
NT: no treatment
OCD: obsessive compulsive disorder
OQ: Outcome Questionnaire
PACQ: Panic Cognitions Questionnaire
PAI: Panic Appraisal Inventory
PARS: Phobic Avoidance Rating Scale
PAS: Panic Agoraphobia Scale/Personality Assessment Schedule
PASQ: Panic Symptoms Questionnaire
PBQ: Panic Belief Questionnaire
PCT: panic control therapy
PD/A: Panic disorder with or without agoraphobia
PDS-MI: Panic Disorder Scale and Mobility Inventory
PDSS: Panic Disorder Severity Scale

PE: psychoeducation
PFPP: panic focused psychodynamic therapy
PPGAS: Panic, Phobia and Generalized Anxiety Scale
PQ: Presence Questionnaire
PSEQ: Panic Self-Efficacy Questionnaire
PSWQ: Penn State Worry Questionnaire
PT: physiological therapies
PTSD: post-traumatic stress disorder
QOLI: Quality of Life Inventory
RR: respiration rate
SAS-SR: Social Adjustment Scale-Self-Report
SCL: Symptom Check List
SD: standard deviation
SE: standard error
SESA: Self-Efficacy Scales for Agoraphobia
SP: supportive psychotherapy
SPRAS: Sheehan Patient-Rated Anxiety Scale
SSQ: Simulation Sickness Questionnaire
ST: short-term
STAI: State–Trait Anxiety Inventory
SUD: Subjective Units of Discomfort
TAU: treatment as usual
VLQ: Valued Living Questionnaire
VRET: virtual reality exposure therapy
WBSI: White Bear Suppression Inventory
WL: wait list
WSA: Work and Social Adjustment scale

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Andersson 2011	Mixed sample. More than 20% of patients are under 18 (personal communication) [Ongoing study]
Barlow 2000a	Drug versus placebo versus CBT versus CBT + drug versus CBT + placebo. In the related study exploring EFT versus CBT, not all patients in the EFT arm were randomised.
Benecke 2014	All participants are required to have a comorbid personality disorder [Ongoing study]
Borden 1986	Assignment to WL was not randomised
Bélanger 2006	Quasi-randomised design (personal communication)
Elsesser 2002	Quasi-randomised design (personal communication)
Fava 1997	The study focuses on refractory patients
Gloster 2010a	The study focuses on refractory patients
Ito 2001	Since data for the intervention arms include data from re-randomised WL patients, data for WL arm cannot be extracted (double-count): the 3 remaining arms would be BT(I) versus BT(E) versus BT(I+E). Comparison not feasible. Randomisation probably not respected: *"twenty patients left the trial before week 4 and were replaced"*, *"the final sample of 80 patients included 10 of the WL who were re-randomised"*. Number of patients originally randomised to each arm unclear. Assessment of the original arms (without the re-randomised and the substitutes) not reported.
Michelson 1996	Replacements are not evenly distributed and constitute more than 15% of the final sample

Study	Reason for exclusion
Teusch 1996	Replacements constitute 24.5% of the final sample
Zane 1993	The study explicitly focuses on agoraphobia, reporting that only 73% of participants had a comorbid panic disorder. We considered studies focusing on agoraphobia, rather than panic disorder, when it could be safely assumed that at least 80% of the participants were suffering from panic disorder.

BT: behaviour therapy
CBT: cognitive behaviour therapy
EFT: emotion focused therapy
WL: wait list

Characteristics of studies awaiting assessment *[ordered by study ID]*

Bressi 2010a

Methods	Randomised controlled trial
Participants	35 patients with a diagnosis of panic disorder (DSM IV-TR)
Interventions	CBT versus PD versus NT (usual care)
Outcomes	• Hamilton Rating Scale for Depression (HAM-D) • Hamilton Rating Scale for Anxiety (HAM-A) • Panic and Anticipatory Anxiety Scale (PAAAS) • Toronto Alexithymia Scale–20 items (TAS-20)
Notes	Only abstract available. No reply after trying to contact the author at the available email addresses.

Foley 2006

Methods	Randomised controlled trial
Participants	Mixed sample (diagnosis for depressive and/or anxiety related disorders according the Anxiety Disorders Interview Schedule)
Interventions	CBT versus 3W (each administered in 8 weekly group therapy sessions)
Outcomes	• Hamilton Depression Rating Scale • Depression and Anxiety Stress Scales • Freiburg Mindfulness Inventory
Notes	Only abstract available. No reply after trying to contact the author at the available email addresses.

Franklin 1990

Methods	Unclear
Participants	56 patients with chronic panic disorder

Franklin 1990 *(Continued)*

Interventions	BT + SMT (self mastery training) versus BT + APP (imaginal rehearsal)
Outcomes	• Anxiety • Panic frequency • Phobic avoidance • Help seeking • Drug usage • Composite criterion of clinical improvement
Notes	Only abstract available. No reply after trying to contact the author at the available email addresses.

Irgens 2009

Methods	Allocation: randomised Endpoint classification: efficacy study Intervention model: single group assignment Masking: single-blind (outcomes assessor) Primary purpose: treatment
Participants	72 patients, age 18 years or older, diagnosis of agoraphobia, score on Mobility Inventory "Alone" of 2.5 or more
Interventions	TFT (thought field therapy: not included) versus CBT versus WL
Outcomes	Change in agoraphobic situation scores in ADIS. Among secondary outcomes: MI, ACQ, BSQ, BDI, BAI.
Notes	(Personal communication)

Margraf 1991

Methods	Unclear
Participants	Unclear
Interventions	Unclear
Outcomes	Unclear
Notes	Full report not retrievable. No reply after trying to contact the author at the available email addresses.

Milrod 2006b

Methods	Allocation: randomised Endpoint classification: efficacy study Intervention model: parallel assignment Masking: single-blind (outcomes assessor) Primary purpose: treatment

Milrod 2006b *(Continued)*

Participants	Age 18 to 70 years, DSM-IV diagnosis criteria for primary PD with or without agoraphobia, history of at least 1 spontaneous panic attack per week within the month prior to study entry
Interventions	CBT versus PD versus PT
Outcomes	Panic Disorder Severity Scale (PDSS)
Notes	Only protocol available. Study completed but not yet published.

Richards 1997

Methods	Unclear
Participants	Unclear
Interventions	Unclear
Outcomes	Unclear
Notes	Full report not retrievable. No reply after trying to contact the author at the available email addresses.

Roache 1998

Methods	Randomised controlled trial
Participants	30 patients with generalised anxiety or panic disorder
Interventions	CBT versus NT
Outcomes	• Self administration of capsules (alprazolam/placebo) • Level of anxiety (scale used not specified)
Notes	Only abstract. No reply after trying to contact the author at the available email addresses.

Strauss 1997

Methods	Unclear
Participants	58 patients with the DSM-III-R diagnosis of panic disorder with and without agoraphobia
Interventions	CBT versus PT versus CBT+Imipramine versus BT + imipramine
Outcomes	• Beck Anxiety Inventory • Fear Questionnaire • Fear Diary • HAM-D • HAM-A • Global Improvement Scale

Strauss 1997 *(Continued)*

	• Behaviour test (DBTA)
Notes	Only abstract. No reply after trying to contact the author at the available email addresses.

Vincelli 2003

Methods	Randomised controlled trial
Participants	12 patients with a *"DSM-IV diagnosis of anxiety disorders for a minimum of 6 months as determined by independent clinicians on clinical interviews."* Exclusion criteria are: psychotic or bipolar disorders, high suicidal risks, medical illness (i.e. cardiac conduction disease, vestibular dysfunction), pregnant women
Interventions	Experiential-cognitive therapy (ECT) versus CBT versus WL
Outcomes	• Beck Depression Inventory (BDI) • State-Trait Anxiety Inventory (STAI) • Agoraphobic Cognitions Questionnaire (ACQ) • Fear Questionnaire (FQ)
Notes	The diagnosis of participants is unclear. No reply after trying to contact the author at the available email addresses.

Vincelli 2004

Methods	Randomised controlled trial
Participants	36 patients (age 35 to 53) with a diagnosis of panic disorder with agoraphobia
Interventions	CBT (12 sessions) versus CBT plus virtual reality exposure (8 sessions) versus WL
Outcomes	• Panic frequency • Level of depression (scale not specified) • State and trait anxiety (scale not specified)
Notes	Only abstract. Maybe continuation of Vincelli 2003a. No reply after trying to contact the author at the available email addresses.

3W: third-wave
ACQ: Agoraphobic Cognitions Questionnaire
ADIS: Anxiety Disorders Interview Schedule
APP: attention/psychological placebo
BAI: Beck Anxiety Inventory
BDI: Beck Depression Inventory
BSQ: Body Sensations Questionnaire
BT: behaviour therapy
CBT: cognitive behaviour therapy
DSM: Diagnostic and Statistical Manual of Mental Disorders
FQ: Fear Questionnaire
HAM-A: Hamilton Anxiety Rating Scale
HAM-D: Hamilton Depression Scale
PAAAS: Panic and Anticipatory Anxiety Scale
PD: psychodynamic therapies

PDSS: Panic Disorder Severity Scale
PT: physiological therapies
MI: Mobility Inventory
NT: no treatment
SMT: self mastery training
STAI: State-Trait Anxiety Inventory
WL: wait list

Characteristics of ongoing studies *[ordered by study ID]*

Barlow 2010

Trial name or title	NCT01243606
Methods	Allocation: randomised Endpoint classification: efficacy study Intervention model: parallel assignment Masking: open-label Primary purpose: treatment
Participants	Age 18 years or older, fluent in the English language, who have a principal DSM-IV diagnosis of SAD, PD/A, GAD or OCD
Interventions	CBT (disorder specific) versus CBT (unified protocol) versus WL
Outcomes	• Anxiety Disorders Interview Schedule for DSM-IV-Lifetime Version (ADIS-IV) • Clinical Global Impression Severity (CGI-S) and Improvement Scales (CGI-I) • Structured Interview Guide for the Hamilton Anxiety and Depression Rating Scale (SIGH-A and SIGH-D)
Starting date	December 2010
Contact information	David H Barlow: dhbarlow@bu.edu Todd J Farchione: tfarchio@bu.edu
Notes	Completion expected by December 2014

Caspi 2012

Trial name or title	NCT01677429
Methods	Allocation: randomised Endpoint classification: efficacy study Intervention model: parallel assignment Masking: double-blind (subject, caregiver, outcomes assessor) Primary purpose: treatment
Participants	Age 18 to 45 years, clinical diagnosis of panic disorder, stable on the same drug and dosage for at least 1 month
Interventions	BT (with VR balance challenge) versus BT (without VR balance challenge) versus CBT
Outcomes	• Hamilton Rating Scale for Anxiety (HAM-A) • Panic Disorder Severity Scale (PDSS) • Panic and Agoraphobia Scale (PAS)

Caspi 2012 *(Continued)*

Starting date	September 2012
Contact information	Revital Amiaz: amiazr@gmail.com Efrat Czerniak: efrat30.3@gmail.com
Notes	Completion expected by September 2013

Gensichen 2012

Trial name or title	DRKS00004386
Methods	Study type: interventional Allocation: randomised controlled trial Blinding: open (masking not used) Control: active control Purpose: treatment Assignment: parallel
Participants	Age 18 or older, clinical diagnosis of panic disorder with or without agoraphobia (ICD-10: F.41.0 or F40.01), positive screening questionnaires, sufficient German language skills, private telephone
Interventions	CBT versus NT
Outcomes	Severity of anxiety, measured by the Beck Anxiety Inventory (BAI)
Starting date	October 2012
Contact information	Thomas Hiller: Thomas.Hiller@med.uni-jena.de
Notes	Completion expected by Spring 2015

Sandell 2012

Trial name or title	NCT01606592
Methods	Allocation: randomised Endpoint classification: efficacy study Intervention model: factorial assignment Masking: open-label Primary purpose: treatment
Participants	Age between 18 and 60, DSM-V diagnosis of panic disorder (with or without agoraphobia), willingness to stop other ongoing psychotherapy treatments and to refrain from non-study treatments during follow-up
Interventions	Randomised CBT/PD versus chosen CBT/PD versus WL
Outcomes	• Panic Disorder Severity Scale (PDSS) • Occupational status • Absence from work due to sickness
Starting date	January 2010

Sandell 2012 *(Continued)*

Contact information	Not specified
Notes	Completion expected by 2017

Teismann 2012

Trial name or title	NCT01680237
Methods	Allocation: randomised Endpoint classification: efficacy study Intervention model: parallel assignment Masking: open-label Primary purpose: treatment
Participants	Age between 18 and 65 years, diagnosis of panic disorder with agoraphobia
Interventions	CBT versus BT
Outcomes	Change (from baseline) in the Mobility Inventory
Starting date	October 2011
Contact information	Tobias Teismann: tobias.teismann@rub.de Juergen Margraf: juergen.margraf@rub.de
Notes	Completion expected by August 2015

ADIS: Anxiety Disorders Interview Schedule
BAI: Beck Anxiety Inventory
BT: behaviour therapy
CBT: cognitive behaviour therapy
CGI: Clinical Global Impression
DSM: Diagnostic and Statistical Manual of Mental Disorders
GAD: generalised anxiety disorder
HAM-A: Hamilton Rating Scale for Anxiety
ICD-10: International Classification of Diseases
NT: no treatment
OC: obsessive compulsive disorder
PAS: Panic and Agoraphobia Scale
PD: psychodynamic therapies
PD/A panic disorder with/without agoraphobia
PDSS: Panic Disorder Severity Scale
SAD: social anxiety disorder
SIGH-A/SIGH-D: Structured Interview Guide for the Hamilton Anxiety and Depression Rating Scale
VR: virtual reality
WL: wait list

ADDITIONAL TABLES

Table 1. Short-term remission: pairwise meta-analyses and NMA results

Comparison (X vs Y)	Pairwise meta-analyses (common τ = 0.69)				Standard NMA (τ = 0.64)			NMA adjusted for SSE (τ = 0.59)		
	# of studies	OR	CI (lower)	CI (upper)	OR	CI (lower)	CI (upper)	OR	CrI (lower)	CrI (upper)
PT vs SP	-	-	-	-	0.36	0.11	1.18	0.35	0.10	1.11
BT vs SP	-	-	-	-	0.37	0.12	1.15	0.38	0.12	1.19
BT vs PT	2	1.11	0.21	6.00	1.03	0.42	2.28	1.10	0.51	2.50
CT vs SP	-	-	-	-	0.47	0.11	2.07	0.44	0.10	1.90
CT vs PT	1	1.22	0.20	7.48	1.33	0.44	4.05	1.27	0.41	3.91
CT vs BT	1	0.95	0.12	7.47	1.29	0.41	4.07	1.15	0.37	3.61
CBT vs SP	3	0.68	0.24	1.91	0.68	0.25	1.83	0.67	0.25	1.82
CBT vs PT	4	1.56	0.62	3.94	1.90	0.98	3.69	**1.95**	1.02	3.97
CBT vs BT	10	**2.09**	1.10	3.97	**1.84**	1.06	3.22	**1.77**	1.02	3.11
CBT vs CT	1	1.69	0.21	13.47	1.43	0.48	4.23	1.53	0.52	4.68
WL vs SP	-	-	-	-	**0.08**	0.03	0.26	**0.22**	0.06	0.78
WL vs PT	4	**0.21**	0.06	0.70	**0.23**	0.12	0.48	0.64	0.27	1.65
WL vs BT	3	**0.12**	0.04	0.43	**0.23**	0.12	0.45	0.58	0.25	1.36
WL vs CT	2	**0.12**	0.02	0.61	**0.18**	0.06	0.52	0.50	0.15	1.77
WL vs CBT	18	**0.13**	0.07	0.22	**0.12**	0.07	0.21	**0.33**	0.16	0.69
NT vs SP	-	-	-	-	0.26	0.04	1.82	0.24	0.03	1.67
NT vs PT	1	0.70	0.11	4.33	0.73	0.14	3.68	0.70	0.14	3.48

Table 1. Short-term remission: pairwise meta-analyses and NMA results *(Continued)*

NT vs BT	-	-	-	-	0.71	0.13	3.96	0.64	0.11	3.51
NT vs CT	1	0.57	0.09	3.48	0.55	0.11	2.73	0.55	0.11	2.71
NT vs CBT	-	-	-	-	0.38	0.07	2.05	0.36	0.07	1.85
NT vs WL	-	-	-	-	3.11	0.58	16.78	1.10	0.18	6.26
PD vs SP	-	-	-	-	0.71	0.14	3.52	0.71	0.14	3.42
PD vs PT	1	4.21	0.70	25.49	1.99	0.57	7.02	2.05	0.60	7.38
PD vs BT	-	-	-	-	1.94	0.50	7.48	1.88	0.48	7.17
PD vs CT	-	-	-	-	1.50	0.30	7.43	1.61	0.33	8.17
PD vs CBT	1	0.51	0.09	2.99	1.05	0.30	3.68	1.06	0.29	3.66
PD vs WL	-	-	-	-	**8.50**	2.27	31.84	3.21	0.75	12.87
PD vs NT	-	-	-	-	2.73	0.36	20.56	2.91	0.40	22.56

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Note that in the main text, where necessary, we inverted the values presented in Table 4 for an easier presentation, in which an OR greater than 1 stands for a higher number of events (short-term remissions) in the intervention group when compared to the control group. Statistically significant results are written in bold.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CrI: credible interval
CT: cognitive training
NMA: network meta-analysis
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
SSE: small study effects
WL: wait list

Table 2. Short-term remission: I^2 values and their 95% confidence intervals

Comparison	I^2(%)	95% CI
CBT vs SP	0	0 to 90
WL vs CBT	58	29 to 75
CBT vs BT	5	0 to 64
WL vs BT	34	0 to 78
WL vs PT	56	0 to 85
CBT vs PT	0	0 to 85

This values refer to standard meta-analyses, where each comparison has its own heterogeneity variance.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CT: cognitive training
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 3. Short-term remission: ranking of treatments

Treatment	SUCRA
SP	88
CBT	76
PD	73
CT	50
BT	41
PT	35
NT	25
WL	13

BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 4. Short-term response: pairwise meta-analyses and NMA results

Comparison (X vs Y)	Pairwise meta-analyses (common τ = 0.55)				Standard NMA (τ = 0.63)			NMA adjusted for SSE (τ = 0.65)		
	# of studies	OR	CI (lower)	CI (upper)	OR	CI (lower)	CI (upper)	OR	CrI (lower)	CrI (upper)
PT vs SP	-	-	-	-	0.68	0.20	2.31	0.69	0.20	2.41
BT vs SP	-	-	-	-	0.589	0.18	1.93	0.69	0.21	2.33
BT vs PT	2	0.90	0.21	3.90	0.87	0.41	1.88	1.00	0.46	2.23
CT vs SP	-	-	-	-	0.39	0.07	2.07	0.36	0.07	1.96
CT vs PT	1	0.94	0.18	4.97	0.58	0.16	2.09	0.53	0.14	2.04
CT vs BT	1	0.20	0.03	1.46	0.66	0.18	2.51	0.53	0.13	2.07
CBT vs SP	3	1.02	0.38	2.73	1.04	0.36	2.99	1.12	0.40	3.26
CBT vs PT	4	1.23	0.56	2.71	1.54	0.81	2.94	1.62	0.84	3.17
CBT vs BT	10	**1.78**	1.00	3.18	**1.77**	1.02	3.04	1.61	0.92	2.86
CBT vs CT	1	2.92	0.43	19.98	2.65	0.73	9.62	3.08	0.81	12.26
WL vs SP	-	-	-	-	**0.17**	0.05	0.53	0.45	0.12	1.93
WL vs PT	4	**0.15**	0.05	0.44	**0.25**	0.12	0.50	0.65	0.24	1.91
WL vs BT	4	**0.32**	0.14	0.73	**0.29**	0.15	0.54	0.65	0.28	1.68
WL vs CT	1	0.31	0.02	4.27	0.43	0.12	1.61	1.24	0.26	6.42
WL vs CBT	17	**0.19**	0.05	0.31	**0.16**	0.10	0.26	0.40	0.18	1.00
NT vs SP	-	-	-	-	0.16	0.02	1.21	0.15	0.02	1.22
NT vs PT	1	0.31	0.06	1.62	0.24	0.05	1.25	0.22	0.04	1.25
NT vs BT	-	-	-	-	0.27	0.05	1.59	0.22	0.03	1.37

Table 4. Short-term response: pairwise meta-analyses and NMA results *(Continued)*

NT vs CT	1	0.33	0.06	1.68	0.41	0.08	2.13	0.42	0.07	2.31
NT vs CBT	-	-	-	-	**0.16**	0.03	0.87	**0.14**	0.02	0.80
NT vs WL	-	-	-	-	0.96	0.17	5.46	0.33	0.04	2.32
PD vs SP	-	-	-	-	1.02	0.20	5.28	1.07	0.20	5.70
PD vs PT	1	4.22	0.84	21.31	1.51	0.42	5.35	1.54	0.42	5.92
PD vs BT	-	-	-	-	1.73	0.44	6.70	1.54	0.37	6.41
PD vs CT	-	-	-	-	2.59	0.45	14.82	2.94	0.47	18.59
PD vs CBT	1	0.34	0.07	1.76	0.98	0.27	3.47	0.95	0.25	3.60
PD vs WL	-	-	-	-	**6.02**	1.60	22.61	2.37	0.49	10.75
PD vs NT	-	-	-	-	6.28	0.80	49.20	7.03	0.84	61.98

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Note that in the main text, where necessary, we inverted the values presented in Table 4 for an easier presentation, in which an OR greater than 1 stands for a higher number of events (short-term remissions) in the intervention group when compared to the control group. Statistically significant results are written in bold.

BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CrI: credible interval
CT: cognitive training
NMA: network meta-analysis
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
SSE: small study effects
WL: wait list

Table 5. Short-term response: I^2 values and their 95% confidence intervals

Comparison	I^2(%)	95% CI
CBT vs SP	7	0 to 90
WL vs CBT	39	0 to 66
CBT vs BT	22	0 to 62
WL vs BT	26	0 to 72
WL vs PT	14	0 to 87
CBT vs PT	45	0 to 82

This values refer to standard meta-analyses, where each comparison has its own heterogeneity variance.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CT: cognitive training
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 6. Short-term response: ranking of treatments

Treatment	SUCRA
CBT	84
PD	74
SP	72
BT	53
PT	52
WL	31
CT	27
NT	7

BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 7. Short-term dropouts: pairwise meta-analyses and NMA results

Comparison	Pairwise meta-analysis				Network meta-analysis (τ = 0.33)		
	# of studies	OR	CI (lower)	CI (upper)	OR	CI (lower)	CI (upper)
PT vs SP	-	-	-	-	0.76	0.28	2.08
BT vs SP	-	-	-	-	0.92	0.35	2.28
BT vs PT	2	2.19	0.46	10.38	1.20	0.60	2.40
CT vs SP	-	-	-	-	1.06	0.20	5.62
CT vs PT	1	5.59	0.59	52.73	1.38	0.33	5.86
CT vs BT	-	-	-	-	1.15	0.26	5.20
CBT vs SP	3	0.63	0.21	1.89	0.64	0.28	1.43
CBT vs PT	3	0.56	0.24	1.28	0.83	0.46	1.50
CBT vs BT	10	0.89	0.52	1.51	0.69	0.42	1.15
CBT vs CT	-	-	-	-	0.60	0.14	2.60
WL vs SP	-	-	-	-	0.48	0.19	1.22
WL vs PT	5	0.74	0.30	1.83	0.63	0.35	1.15
WL vs BT	4	**0.34**	0.16	0.69	**0.52**	0.30	0.93
WL vs CT	-	-	-	-	0.46	0.11	1.97
WL vs CBT	14	0.70	0.42	1.16	0.76	0.48	1.20
NT vs SP	-	-	-	-	0.10	0.00	2.18
NT vs PT	1	0.29	0.01	7.51	0.13	0.01	3.10
NT vs BT	-	-	-	-	0.11	0.00	2.69

Table 7. Short-term dropouts: pairwise meta-analyses and NMA results *(Continued)*

NT vs CT	1	0.07	0.00	1.37	0.10	0.00	1.93
NT vs CBT	-	-	-	-	0.16	0.01	3.80
NT vs WL	-	-	-	-	0.21	0.09	5.01
PD vs SP	-	-	-	-	0.33	0.08	1.46
PD vs PT	1	**0.16**	0.03	0.84	0.44	0.12	1.52
PD vs BT	-	-	-	-	0.36	0.10	1.35
PD vs CT	-	-	-	-	0.32	0.05	2.04
PD vs CBT	1	1.21	0.27	5.35	0.52	0.15	1.80
PD vs WL	-	-	-	-	0.69	0.19	2.50
PD vs NT	-	-	-	-	3.31	0.11	97.03
APP vs SP	-	-	-	-	0.84	0.07	10.61
APP vs PT	-	-	-	-	1.10	0.10	12.30
APP vs BT	-	-	-	-	0.92	0.08	10.49
APP vs CT	1	0.97	0.08	11.51	0.80	0.09	7.21
APP vs CBT	-	-	-	-	1.32	0.12	14.61
APP vs WL	-	-	-	-	1.75	0.16	19.51
APP vs NT	-	-	-	-	8.37	0.21	333.81
APP vs PD	-	-	-	-	2.53	0.17	36.70
PE vs SP	-	-	-	-	0.84	0.01	97.42
PE vs PT	-	-	-	-	1.10	0.01	119.64
PE vs BT	-	-	-	-	0.92	0.01	100.89

Table 7. Short-term dropouts: pairwise meta-analyses and NMA results *(Continued)*

PE vs CT	-	-	-	-	0.80	0.01	77.96
PE vs CBT	-	-	-	-	1.32	0.01	142.93
PE vs WL	-	-	-	-	1.75	0.02	189.75
PE vs NT	-	-	-	-	8.36	0.04	1954.05
PE vs PD	-	-	-	-	2.53	0.02	316.00
PE vs APP	-	-	-	-	1.00	0.02	55.66
3W vs SP	-	-	-	-	0.36	0.06	2.12
3W vs PT	-	-	-	-	0.47	0.08	2.65
3W vs BT	-	-	-	-	0.39	0.07	2.15
3W vs CT	-	-	-	-	0.34	0.04	3.03
3W vs CBT	1	0.56	0.13	2.51	0.56	0.11	2.88
3W vs WL	-	-	-	-	0.74	0.14	4.04
3W vs NT	-	-	-	-	3.55	0.10	126.65
3W vs PD	-	-	-	-	1.07	0.13	8.31
3W vs APP	-	-	-	-	0.42	0.02	7.74
3W vs PE	-	-	-	-	0.42	0.00	60.45

OR (X vs Y) is defined as (Odds X)/(Odds Y). Of course, since this is dropout outcome, an OR smaller than one favours treatment X. Statistically significant results are written in bold.
3W: third-wave
APP: attention/psychological placebo
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CrI: credible interval
CT: cognitive training
NMA: network meta-analysis
NT: no treatment

PD: psychodynamic therapies
PE: psychoeducation
PT: physiological therapies
SP: supportive psychotherapy
SSE: small study effects
WL: wait list

Table 8. Short-term dropouts: I^2 values and their 95% confidence intervals

Comparison	I^2(%)	95% CI
CBT vs SP	49	0 to 85
WL vs CBT	0	0 to 55
CBT vs BT	0	0 to 62
WL vs BT	0	0 to 85
WL vs PT	25	0 to 70
CBT vs PT	12	0 to 91

This values refer to standard meta-analyses, where each comparison has its own heterogeneity variance.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CT: cognitive training
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 9. Short-term dropouts: ranking of treatments

Treatment	SUCRA
NT	85
PD	75
WL	67
3W	67
CBT	51
APP	42
PE	42
PT	38
CT	29
BT	28
SP	26

Note that higher ranking treatments correspond to lower dropout rate.
3W: third-wave
APP: attention/psychological placebo
BT: behaviour therapy
CBT: cognitive behaviour therapy

CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PE: psychoeducation
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 10. Short-term improvement: pairwise meta-analyses and NMA results

Comparison (X vs Y)	Pairwise meta-analysis				Network meta-analysis (τ = 0.34)		
	# of studies	SMD	CI (lower)	CI (upper)	SMD	CI (lower)	CI (upper)
PT vs SP	-	-	-	-	0.24	-0.35	0.84
BT vs SP	-	-	-	-	0.16	-0.42	0.74
BT vs PT	1	-0.26	-0.98	0.46	-0.08	-0.47	0.31
CT vs SP	-	-	-	-	0.17	-0.56	0.90
CT vs PT	1	-0.02	-0.67	0.64	-0.08	-0.63	0.48
CT vs BT	1	0.65	-0.16	1.46	0.00	-0.56	0.57
CBT vs SP	3	-0.08	-0.40	0.24	-0.05	-0.56	0.47
CBT vs PT	5	-0.05	-0.30	0.19	-0.29	-0.60	0.02
CBT vs BT	10	**-0.24**	-0.45	-0.03	-0.21	-0.48	0.07
CBT vs CT	2	0.03	-0.88	0.94	-0.21	-0.73	0.31
WL vs SP	-	-	-	-	**1.05**	0.49	1.60
WL vs PT	4	**0.87**	0.09	1.65	**0.80**	0.47	1.13
WL vs BT	3	**0.92**	0.59	1.26	**0.89**	0.57	1.20
WL vs CT	1	0.60	-0.27	1.47	**0.88**	0.34	1.42
WL vs CBT	17	**1.14**	0.87	1.41	**1.09**	0.88	1.31
NT vs SP	-	-	-	-	0.78	-0.06	1.63
NT vs PT	1	0.36	-0.26	0.98	-0.46	-1.12	0.21
NT vs BT	-	-	-	-	0.62	-0.09	1.33

Table 10. Short-term improvement: pairwise meta-analyses and NMA results *(Continued)*

NT vs CT	1	0.39	-0.25	1.03	0.61	-0.11	1.34
NT vs CBT	1	**1.30**	0.46	2.14	**0.83**	0.16	1.50
NT vs WL	-	-	-	-	-0.27	-0.96	0.42
PD vs SP	-	-	-	-	-0.21	-1.05	0.63
PD vs PT	1	**-1.18**	-1.79	-0.57	-0.46	-1.12	0.21
PD vs BT	-	-	-	-	-0.38	-1.09	0.34
PD vs CT	-	-	-	-	-0.38	-1.21	0.45
PD vs CBT	1	0.57	-0.07	1.2	-0.17	-0.83	0.50
PD vs WL	-	-	-	-	-1.26	-1.95	-0.57
PD vs NT	-	-	-	-	-1.00	-1.91	-0.07
PE vs APP	1	-0.25	-0.87	0.38	-	-	-
3W vs SP	-	-	-	-	0.20	-0.82	1.22
3W vs PT	-	-	-	-	-0.04	-0.98	0.90
3W vs BT	-	-	-	-	0.04	-0.89	0.97
3W vs CT	-	-	-	-	0.04	-0.99	1.07
3W vs CBT	1	0.26	-0.33	0.84	0.25	-0.64	1.14
3W vs WL	-	-	-	-	-0.84	-1.76	0.07
3W vs NT	-	-	-	-	-0.58	-1.69	0.53
3W vs PD	-	-	-	-	0.42	-0.69	1.53

A positive SMD for X vs Y means treatment Y is better than X, the opposite for a negative SMD (this is because for the scales used a lower score corresponds to a better treatment).
Statistically significant results are written in bold.
3W: third-wave
APP: attention/psychological placebo

BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PE: psychoeducation
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 11. Short-term improvement: I^2 values and their 95% confidence intervals

Comparison	I^2(%)	95% CI
CBT vs SP	0	0 to 90
CBT vs PT	0	0 to 79
WL vs PT	79	45 to 92
CBT vs BT	0	0 to 62
WL vs BT	23	0 to 92
WL vs CBT	61	34 to 77

This values refer to standard meta-analyses, where each comparison has its own heterogeneity variance.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CT: cognitive training
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 12. Short-term improvement: ranking of treatments

Treatment	SUCRA
PD	83
CBT	79
SP	69
CT	54
3W	53
BT	52
PT	43
NT	14
WL	4

3W: third-wave
BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 13. Long-term remission/response: pairwise meta-analyses and NMA results

Comparison (X vs Y)	Pairwise meta-analyses				Network meta-analysis		
	# of studies	OR	CI (lower)	CI (upper)	OR	CI (lower)	CI (upper)
PT vs SP	-	-	-	-	1.31	0.33	5.25
BT vs SP	-	-	-	-	1.27	0.38	4.27
BT vs PT	1	1.17	0.28	4.87	0.97	0.37	2.57
CT vs SP	-	-	-	-	0.96	0.21	4.36
CT vs PT	1	0.83	0.25	2.80	0.74	0.26	2.06
CT vs BT	1	0.38	0.08	1.86	0.76	0.28	2.31
CBT vs SP	1	2.09	0.80	5.47	2.09	0.73	5.98
CBT vs PT	2	1.33	0.47	3.76	1.60	0.64	3.95
CBT vs BT	5	1.66	0.80	3.44	1.64	0.90	2.97
CBT vs CT	1	1.56	0.34	7.11	2.16	0.73	6.37
PD vs SP	-	-	-	-	1.67	0.33	8.33
PD vs PT	-	-	-	-	1.28	0.28	5.81
PD vs BT	-	-	-	-	1.31	0.34	5.07
PD vs CT	-	-	-	-	1.73	0.34	8.79
PD vs CBT	1	0.80	0.26	9.86	0.80	0.24	2.69

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Note that in the main text, where necessary, we inverted the values presented in Table 4 for an easier presentation, in which an OR greater than 1 stands for a higher number of events (short-term remissions) in the intervention group when compared to the control group.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
PD: psychodynamic therapies
PT: physiological therapies

SP: supportive psychotherapy
WL: wait list

Table 14. Long-term remission/response: ranking of treatments

Treatment	SUCRA
CBT	85
PD	64
PT	49
BT	43
SP	31
CT	27

BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 15. Sensitivity analyses: pairwise meta-analyses and NMA results for short-term remission excluding group therapy trials

Comparison	Pairwise meta-analyses (common τ 0.05)				Standard NMA (τ ≅ 0)			NMA adjusted for SSE (τ ≅ 0.60)		
	# of studies	OR	CI (lower)	CI (upper)	OR	CI (lower)	CI (upper)	OR	CrI (lower)	CrI (upper)
PT vs SP	-	-	-	-	**0.27**	0.11	0.64	0.34	0.10	1.08
BT vs SP	-	-	-	-	**0.35**	0.16	0.75	0.38	0.12	1.16
BT vs PT	2	0.98	0.26	3.74	1.29	0.69	2.42	1.10	0.51	2.52
CT vs SP	-	-	-	-	0.43	0.13	1.47	0.44	0.10	1.87
CT vs PT	-	-	-	-	1.60	0.51	4.96	1.27	0.41	3.90
CT vs BT	1	0.95	0.19	4.84	1.24	0.44	3.51	1.15	0.37	3.68
CBT vs SP	3	0.67	0.33	1.35	0.67	0.34	1.27	0.67	0.24	1.77
CBT vs PT	3	2.02	0.85	4.76	**2.46**	1.39	4.33	**1.94**	1.02	3.97
CBT vs BT	9	**1.90**	1.15	3.16	**1.90**	1.26	2.88	**1.76**	1.02	3.13
CBT vs CT	1	1.69	0.33	8.76	1.54	0.55	4.32	1.53	0.52	4.66
WL vs SP	-	-	-	-	**0.07**	0.03	0.15	**0.22**	0.06	0.76
WL vs PT	4	**0.22**	0.09	0.53	**0.25**	0.14	0.45	0.64	0.27	1.66
WL vs BT	3	**0.17**	0.08	0.36	**0.19**	0.12	0.31	0.58	0.26	1.36
WL vs CT	2	**0.12**	0.03	0.46	**0.16**	0.06	0.43	0.51	0.15	1.82
WL vs CBT	11	**0.10**	0.05	0.19	**0.10**	0.07	0.16	**0.33**	0.16	0.70
PD vs SP	-	-	-	-	0.61	0.21	1.80	0.24	0.03	1.64
PD vs PT	1	**4.22**	1.17	15.15	2.25	0.93	5.42	0.70	0.14	3.66
PD vs BT	-	-	-	-	1.74	0.68	4.47	0.64	0.11	3.60

Table 15. Sensitivity analyses: pairwise meta-analyses and NMA results for short-term remission excluding group therapy trials *(Continued)*

PD vs CT	-	-	-	-	1.41	0.37	5.34	0.55	0.11	2.74
PD vs CBT	1	0.51	0.15	1.76	0.91	0.38	2.19	0.36	0.07	1.87
PD vs WL	-	-	-	-	**8.97**	3.50	22.95	1.09	0.18	6.45

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Note that in the main text, where necessary, we inverted the values presented in Table 4 for an easier presentation, in which an OR greater than 1 stands for a higher number of events (short-term remissions) in the intervention group when compared to the control group. Statistically significant results are written in bold.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CI: confidence interval
CrI: credible interval
CT: cognitive training
NMA: network meta-analysis
NT: no treatment
PD: psychodynamic therapies
PE: psychoeducation
PT: physiological therapies
SP: supportive psychotherapy
SSE: small study effects
WL: wait list

Table 16. Sensitivity analyses: I^2 values and their 95% confidence intervals for short-term remission excluding group therapy trials

Comparison	I^2(%)	95% CI
WL vs PT	56	0 to 85
WL vs BT	34	0 to 78
WL vs CBT	21	0 to 60
CBT vs BT	6	0 to 67
CBT vs SP	0	0 to 90
CBT vs PT	0	0 to 90

This values refer to standard meta-analyses, where each comparison has its own heterogeneity variance.
BT: behaviour therapy
CBT: cognitive behaviour therapy
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 17. Sensitivity analyses: ranking of treatments for short-term remission excluding group therapy trials

Treatment	SUCRA
SP	89
CBT	76
PD	73
CT	50
BT	40
PT	35
NT	26
WL	13

BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 18. Sensitivity analyses: pairwise meta-analyses and I^2 for short-term remission excluding trials in which a concomitant pharmacotherapy is allowed

Comparison (X vs Y)	Pairwise meta-analyses				
	# of studies	OR	CI (lower)	CI (upper)	I^2
WL vs CBT	3	**0.07**	0.01	0.30	0.0%
CBT vs BT	2	**2.21**	0.96	5.10	0.7%
CBT vs SP	2	0.51	0.22	1.23	0.0%
WL vs BT	1	**0.22**	0.11	0.44	-
NT vs CT	1	0.57	0.17	1.91	-
CT vs PT	1	1.22	0.36	4.12	-
NT vs PT	1	0.70	0.27	1.10	-

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Note that in the main text, where necessary, we inverted the values presented in Table 4 for an easier presentation, in which an OR greater than 1 stands for a higher number of events (short-term remissions) in the intervention group when compared to the control group. Statistically significant results are written in bold.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 19. Sensitivity analyses: pairwise meta-analyses with their I^2 and NMA results for short-term remission excluding from the analyses trials in which pharmacotherapy stabilisation was not required

Comparison (X vs Y)	Pairwise meta-analyses					Network meta-analysis (tau = 0.63)		
	# of studies	OR	CI (lower)	CI (upper)	I^2	OR	CI (lower)	CI (upper)
BT vs PT	-	-	-	-	-	0.56	0.12	2.67
CBT vs PT	3	1.56	0.69	3.55	30	1.80	0.69	4.67
CBT vs BT	2	4.04	0.73	22.52	72	3.16	0.93	10.71
WL vs PT	3	0.22	0.01	4.04	69	**0.11**	0.03	0.36
WL vs BT	1	**0.02**	0.00	0.32	-	**0.19**	0.04	0.87
WL vs CBT	6	**0.06**	0.03	0.014	0	**0.06**	0.02	0.16

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Note that in the main text, where necessary, we inverted the values presented in Table 4 for an easier presentation, in which an OR greater than 1 stands for a higher number of events (short-term remissions) in the intervention group when compared to the control group. Statistically significant results are written in bold.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

Table 20. Sensitivity analyses: pairwise meta-analyses for short-term remission using a fixed-effect rather than a random-effects model

Comparison (X vs Y)	# of studies	Random-effects pairwise meta-analyses			Fixed-effect pairwise meta-analyses		
		OR	CI (lower)	CI (upper)	OR	CI (lower)	CI (upper)
BT vs PT	2	1.09	0.22	5.52	0.96	0.27	3.46
CT vs PT	1	1.22	0.36	4.12	1.22	0.36	4.12
CT vs BT	1	0.95	0.20	4.54	0.95	0.20	4.54

Table 20. Sensitivity analyses: pairwise meta-analyses for short-term remission using a fixed-effect rather than a random-effects model *(Continued)*

CBT vs SP	3	0.67	0.35	1.27	0.67	0.35	1.27
CBT vs PT	4	1.43	0.78	2.62	1.43	0.78	2.62
CBT vs BT	10	**2.05**	1.29	3.27	**2.05**	1.30	3.22
CBT vs CT	1	1.69	0.35	8.22	1.69	0.35	8.22
PT vs WL	4	4.55	0.89	25.00	**4.55**	1.96	10.00
BT vs WL	3	**8.33**	2.33	25.00	**5.56**	2.94	11.11
CT vs WL	2	**8.33**	2.22	33.33	**8.33**	2.22	33.33
CBT vs WL	18	**9.09**	4.76	20.00	**5.56**	3.70	8.33
PT vs NT	1	1.43	0.41	5.00	1.43	0.41	5.00
CT vs NT	1	1.75	0.52	5.88	1.75	0.52	5.88
PD vs PT	1	**4.22**	1.27	13.95	**4.22**	1.27	13.95
PD vs CBT	1	0.51	0.16	1.613	0.51	0.16	1.613

OR (X vs Y) is defined as (Odds X)/(Odds Y). For each comparison X vs Y, an OR greater than one favours treatment X, an OR less than one favours treatment Y. Statistically significant results are written in bold.
BT: behaviour therapy
CBT: cognitive behaviour therapy
CT: cognitive training
NT: no treatment
PD: psychodynamic therapies
PT: physiological therapies
SP: supportive psychotherapy
WL: wait list

APPENDICES

Appendix 1. CCDANCTR-References Register search (psychotherapies for panic)

1. (therap* or psychotherap*) [ti,ab]

2. psychotherapy [kw]

3. (acceptance* or commitment* or “activity scheduling” or adlerian or art or aversion or brief or “client cent*” or cognitive* or color or colour or compassion-focused or “compassion* focus*” or compassionate or conjoint or conversion or conversational or couples or dance or dialectic* or diffusion or distraction or eclectic or (emotion and focus*) or emotion-focus* or existential or experiential or exposure or expressive or family or focus-oriented or “focus oriented” or freudian or gestalt or “group” or humanistic or implosive or insight or integrative or interpersonal or jungian or kleinian or logo or marital or metacognitive or meta-cognitive or milieu or morita or multimodal or multi-modal or music or narrative or nondirective or non-directive or “non directive” or nonspecific or non-specific or “non specific” or “object relations” or “personal construct” or “person cent*” or person-cent* or persuasion or play or ((pleasant or pleasing) and event*) or primal or problem-focused or “problem focused” or problem-solving or “problem solving” or process-experiential or “process experiential” or psychodynamic or “rational emotive” or reality or “reciprocal inhibition” or relationship* or reminiscence or restructuring or rogerian or schema* or self-control* or “self control*” or “short term” or short-term or sex or “social effectiveness” or “social skill*” or socio-environment* or “socio environment*” or “solution focused” or solution-focused or “stress management” or supportive or time-limited or “time limited” or “third wave” or transference or transtheoretical or validation)

4. (abreaction or “acting out” or “age regression” or ((assertive* or attention or autogenic or mind or sensitivity) and train*) or autosuggestion or “balint group” or ((behavior* or behaviour*) and (activation or therap* or treatment or contracting or modification)) or bibliotherap* or biofeedback or catharsis or *cognitive* or *CBT* or “mind training” or counsel* or “contingency management” or countertransference or “covert sensitization” or “eye movement desensiti*” or EMDR or “crisis intervention” or “dream analysis” or “emotional freedom” or “free association” or “functional analys*” or griefwork or hypno* or imagery or meditation* or “mental healing” or mindfulness* or “panic program” or psychoanaly* or psychodrama or psychoeducat* or (psycho* and support*) or psychotherap* or relaxation or “role play*” or “self analysis” or “self esteem” or “self-help or “self help” or “sensitivity training” or “support group*” or therapist or “therapeutic technique*” or “transactional analysis”)

5. ((1 or 2) and 3) or 4

6. panic

7. (5 and 6)

Appendix 2. PubMed search strategy

((("randomized controlled trial"[Publication Type]) OR ("controlled clinical trial"[Publication Type]) OR ("clinical trials as topic"[MeSH Terms]) OR ((randomized[Title/Abstract]) OR randomised[Title/Abstract]) OR (randomly[Title/Abstract]) OR (placebo[Title/Abstract]) OR (trial[Title])) NOT ("animals"[MeSH Terms] NOT "humans"[MeSH Terms])) AND (("psychotherapy"[MeSH Terms]) OR (psychotherap* OR psychoanaly* OR psychodynamic OR psychodrama OR psychoeducat*[Title/Abstract])) AND (("agoraphobia"[MeSH Terms]) OR ("panic disorder"[MeSH Terms]) OR ("panic"[MeSH]) OR (panic OR agoraphobi*[Title/Abstract]))

Appendix 3. Statistical software details

We used Stata for most analyses presented in this review. We employed the mvmeta command for the network meta-analyses (White 2011). We performed network plots and inconsistency plots using the network_graphs package in Stata (Chaimani 2013).

We used OpenBUGS to fit the small study effects network meta-analyses. We assumed minimally informative prior distributions for the logarithm of the treatment effects, and the heterogeneity standard deviation, $\tau \sim U(0,5)$. The results presented pertain to a sample obtained from two independent chains with 100,000 iterations each, after a 10,000 burn-in period. We confirmed convergence using the Brooks-Gelman-Rubin criterion (Brooks 1998).

CONTRIBUTIONS OF AUTHORS

TAF and AP conceived the study.

AP led the project under TAF supervision.

AP, TAF and GS worked on the protocol; HI, AT and OE provided suggestions and input.

GS and OE performed the statistical analyses.

AP and TAF worked on the review; HI, AT, OE and GS provided suggestions and input.

All authors reviewed and approved the final version of the review.

DECLARATIONS OF INTEREST

AP has no competing interests.

TAF has received honoraria for speaking at CME meetings sponsored by Asahi Kasei, Eli Lilly, GlaxoSmithKline, Mochida, MSD, Otsuka, Pfizer, Shionogi and Tanabe-Mitsubishi. He is a diplomate of the Academy of Cognitive Therapy. He has received royalties from Igaku-Shoin, Seiwa-Shoten and Nihon Bunka Kagaku-sha. He is on the advisory board for Sekisui Chemicals and Takeda Science Foundation. The Japanese Ministry of Education, Science, and Technology, the Japanese Ministry of Health, Labor and Welfare, and the Japan Foundation for Neuroscience and Mental Health have funded his research projects.

HI has no competing interests.

AT has received honoraria for speaking at a meeting sponsored by Eli Lilly.

OE has no competing interests.

GS has no competing interests.

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

We first planned to use scale endpoint data, which typically cannot have negative values and are easier to interpret from a clinical point of view. However, as a post hoc decision, we decided to use change data in an attempt to reduce the amount of heterogeneity due to the baseline imbalance found across studies.

INDEX TERMS

Medical Subject Headings (MeSH)

Agoraphobia [psychology] [*therapy]; Panic Disorder [psychology] [*therapy]; Psychotherapy [*methods]; Randomized Controlled Trials as Topic

MeSH check words

Adult; Humans